Mapping and Phenotyping Genes in the Del(13)Svea36H Region of Mouse Chromosome 13

Thesis

Richard McKeone

Mapping and Phenotyping Genes in the Del(13)Svea36H Region of Mouse Chromosome 13

PhD Thesis

MRC Harwell

Mammalian Genetics Unit

Medical Research Council

Harwell

Didcot

Oxfordshire

OX11-0RD

Date of Submission: 31/10/08

Date of award: 24/04/09

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Acknowledgements

I would like pay tribute to my director of studies, Paul Denny for his enthusiasm and encouragement throughout the course of the project. I would also like to thank my supervisors Ruth Arkell and Steve Brown for their ideas and guidance. Pat Nolan and Cheryl Gregory-Evans provided valuable help in the dissertation write-up. Pat also allowed the use of his project licence for the secondary stages of neurological phenotyping. The preceding work of Cheryl and Helena Vieira was critical for guiding the direction of the eye phenotyping aspect of this study.

Many people were essential for the completion of the laboratory work. Matthew Cadman in particular was very generous in passing on his physical mapping knowledge and skills. Thanks also to Nick Parkinson for his advice regarding the available software and laboratory based tools for functional genetics; Philomena Mburu for advice on cloning methods and protein assays; Andy Haynes for his knowledge of cell culture; and Alison Hough and Deen Quwailid for their advice on dHPLC.

M. Cadman and T. Marsland determined the informativeness of the D13Mit markers. P. Denny and M. Cadman selected the markers used in probing the RPCI-23 BAC libraries. The BAC library screening was performed with C. Sellick, who also carried out T7 and Sp6 BAC end sequencing. M. Cadman analysed the BAC hybridisation data to construct the Del36H BAC contig, he also identified and sequenced the mud91 mutation at Sox4. J. Davies carried out genotyping on all

individuals with marker D13Mit17 before marker panel genotyping was introduced. Ghazala Mirza, Louise Jane Glenny and Kulvinder Kaur of the Genome Core at the Wellcome Trust Centre for Human Genetics carried out marker panel genotyping for 6446 of the 7771 individual animals from all lines except line 91. C. Willoughby and D. Bogani carried out the analysis of the genotyping data for mapping of the critical regions for all lines except line 91. Louise Jane Glenny and Kulvinder Kaur of the Wellcome trust identified the deletion breakpoints.

Following enucleation & fixation, Foxf2^{W174R} eyes and heads were sent to the histology department at Imperial College for sectioning and H&E staining. Cheryl Gregory Evans performed the analysis of the Foxf2^{W174R} eyes. The histology department at MRC Harwell sectioned heads for palate analysis and the lungs. Jim Humphreys prepared samples that were analysed by Michael Cheeseman for necropsy analysis. Hilda Tatteosian assisted with and taught techniques for palate analysis. Sue Morse & Rachel Hardisty assisted with and taught techniques for embryo extractions. Jacky King advised on gel techniques. Charlotte Dean provided assistance with and taught techniques for the embryonic lung extractions and analysis. Lucie Vizor, Sara Wells and the animal technicians at 371 and the Mary Lyon Centre were responsible for animal care and husbandry. Neil Dear and Pat Nolan allowed the use of their project licences for the preliminary phenotype testing and secondary PPI/startle testing respectively. The Phenotyping core performed the SHIRPA, blood collection and initial IPGTT. Tertius Hough performed the blood biochemistry analysis. M. Goldsworthy & L. Moir performed the secondary IPGTT.

I would like to give special thanks to my wife, Ainslie, without whom I would have never thought of beginning a PhD- never mind finishing one! Ainslie and more recently my two sons, Calum and Joshua have kept me laughing through the later stages of the project.

Figures and Tables

FIGURE 1.1: WHOLE GENOME & DELETION MUTAGENESIS SCREENS 6
FIGURE 1.2: CRE-LOX SYSTEM FOR MUTATION INDUCTION 8
FIGURE 1.3: FORKHEAD PROTEINS BINDING TO DNA 25
TABLE 1.1: PHENOTYPES OF MUTATIONS IN VERTEBRATE FORKHEAD-BOX GENES. 33
TABLE 1.2: MARKERS USED BY GOULD ET AL TO CHARACTERISE A 6P TRANSLOCATION 39
FIGURE 1.4: TRANSILLUMINATION DEFECT OF THE HUMAN EYE 41
FIGURE 1.5: PERIOCULAR MESENCHYME IN A MOUSE EYE AT E12.5 41
TABLE 2.1: RESTRICTION ENZYMES USED. 68
TABLE 2.2 PRIMERS USED FOR THE NED SIZE STANDARD USED FOR GENOTYPING 72
TABLE 2.3 DYE LABELS AND VOLUMES USED 74
TABLE 3.1: POOLED HYBRIDISATION PROBES 111
TABLE 3.2: SINGLE HYBRIDISATION PROBES 112
FIGURE 3.1: PYHSICAL MAP OF THE DEL36H REGION 114
FIGURE 3.2: BREEDING SCHEME FOR THE LETHAL RECESSIVE SCREEN 115
FIGURE 3.3: AN SSCP GEL 118
TABLE 3.3: POLYMORPHISMS IDENTIFIED DURING THE MICROSATELLITE SCREEN 119
TABLE 3.4: POLYMORPHISM DATA FOR INFORMATIVE BACS 124
TABLE 3.5: REDUNDANT MARKERS 127
TABLE 3.6: SINGLE NUCLEOTIDE POLYMORPHISMS (SNPS) 129
FIGURE 3.4: POSITION OF GENOTYPING MARKERS 133
TABLE 3.7: MARKER PANELS USED IN THE DIFFERENT GENOTYPING SCHEMES 133
TABLE 3.8: ALTERNATIVE REVERSE PRIMERS FOR GENOTYPING PRIMERS 138
FIGURE 3.5: EXAMPLE OF DEL36H GENOTYPE DETERMINATION 140
FIGURE 3.6: BREEDING SCHEMES FOR INHERITANCE TESTING 145
FIGURE 3.7: CRITICAL REGIONS OF THE DEL 36H MUTANTS 147
TABLE 3.9: GENOTYPING OF MUD91 FOR CONFIRMATION OF SOX4 LOCALISATION 148
FIGURE 4.1: HETERODUPLEXES IN DHPLC MUTATION DETECTION 158
FIGURE 4.2: OPTIMAL DETECTION LIMITS OF THE FOXF2_1_1 FRAGMENT 159
FIGURE 4.3: RELAXATION OF DHPLC HELICAL FRACTION LIMITS 160
FIGURE 4.4 TEMPERATURE AFFECTS MUTATION DETECTION 161
FIGURE 4.5: RELAXATION OF DHPLC TEMPERATURE LIMITS 162
FIGURE 4.6: OPTIMAL DETECTION ZONES OF FOXF2_1_1 AT THREE DHPLC TEMPERATURES 162
TABLE 4.1: PRIMERS USED TO SCREEN THE ARCHIVE. 163
FIGURE 4.7: FOXF2 TRANSCRIPT 164
FIGURE 4.8: HELICAL FRACTIONS FOR EACH AMPLICON 166
FIGURE 4.9: DHPLC ELECTROPHEROGRAMS 167
FIGURE 4.10: ELECTROPHEROGRAMS FOR SPLICE MUTANT 168
FIGURE 4.11: ELECTROPHEROGRAMS FOR MISSENSE MUTATIONS 169
FIGURE 4.12: ELECTROPHEROGRAMS FOR SILENT MUTATIONS 170
FIGURE 4.13: SFCI MUTATION CONFIRMATION OF W174R MUTATION 173
FIGURE 4.14: RESTRICTION SITES IN FOXF2 W174R MUTANT 174
FIGURE 4.15: MUTATION SEQUENCING 176
FIGURE 4.16: FOXF2 MUTATIONS IN RELATION TO THE TRANSCRIPT 177
FIGURE 4.17: FOXF2 MUTATIONS IN RELATION TO THE FUNCTIONAL DOMAINS 178
FIGURE 4.18: SPLICE SITE MUTATION IN RELATION TO THE EXON/INTRON BOUNDARY 182
FIGURE4.19: V412F MUTATION IN RELATION TO THE DOMAIN ARCHITECTURE 183
FIGURE 4.20: W174R MUTATION IN RELATION TO THE FORKHEAD DOMAIN 184
FIGURE 4.21: W174R MUTATION & 3D STRUCTURE OF A RELATED FORKHEAD DOMAIN 187
FIGURE 4.22: EFFECT OF W174R ON SECONDARY STRUCTURE PREDICTIONS 188
FIGURE 4.23: RELATIVE PCR PEAK QUALITY DURING ARCHIVE SCREENING 190
FIGURE 5.1: GENE DRIVEN SCREEN BREEDING SCHEME 198
TABLE 5.1: BEHAVIOURAL TESTING 202
TABLE 5.2: PHYSICAL TRAITS 203
TABLE 4.3: PRIMARY SCREEN STATISTICS 204

TABLE 5.4: LIGHT-DARK BOX TESTING 207
FIGURE 5.2: STARTLE RESPONSE TO STARTLE-2 212
FIGURE 5.3 STARTLE DATA FOR STARTLE-2 PROGRAMME 214
FIGURE 5.4: HISTOGRAMS FOR STARTLE-2 RESPONSE 216
FIGURE 5.5: BOX-PLOT OF STARTLE-2 DATA 217
FIGURE 5.5: PPI-1 RESPONSE 221
FIGURE 5.6: PPI-2 RESPONSE 223
TABLE 5.5: METABOLIC SCREEN 229
FIGURE 5.8: SIGNIFICANT RESULTS FROM METABOLIC SCREEN 230
TABLE 5.6 PROTEIN AND ASCORBIC ACID DIPSTICK UNIT CONVERSION 232
TABLE 5.7: ASCORBIC ACID AND PROTEIN FREQUENCIES 232
FIGURE 5.9: IPGTT PRELIMINARY TESTING 236
FIGURE 5.10: IPGTT SECONDARY TESTING 237
FIGURE 5.11: HETEROZYGOTE WEIGHTS 239
FIGURE 5.12: EYE DEVELOPMENT 241
FIGURE 5.13: ANTERIOR SEGMENT DYSGENESIS CAN CAUSE AN INCREASE IN INTRAOCULAR PRESSURE 242
FIGURE 5.14: IRIS STROMA MALFORMATION 244
FIGURE 5.15: IRIDOCORNEAL ANGLE MALFORMATION 246
FIGURE 5.16: RETINAL AND CORNEAL SECTIONS FROM FOXF2W174R HETEROZYGOTES. 248
FIGURE 5.17: BULGING EYES 249
FIGURE 5.18: CILIARY BODY 250
FIGURE 5.19: FOXF2^{W174R} HETEROZYGOTE OPTIC NERVES 251
FIGURE 5.20: FORMATIVE CILIARY BODY 253
TABLE 5.8: LIVE EMBRYO FREQUENCIES 256
TABLE 5.9: FREQUENCIES OF ALL LIVE BIRTHS IN THE MARY LYON CENTRE. 256
FIGURE 5.21: HOMOZYGOUS PUP 257
FIGURE 5.22: SECONDARY PALATE OF FOXF2^{W174R} HOMOZYGOTES 262
FIGURE 6.1: FOXF2 MOLECULAR INTERACTIONS 268
FIGURE 6.2 EMSA & TRANSACTIVATION ASSAY SCHEMES 275
FIGURE 6.3: INTENDED FOXF2 EXPRESSION PLASMID DESIGN 277
FIGURE 6.4: RNA PURITIES 278
TABLE 6.1: EXPECTED AND OBSERVED PRODUCT SIZES FROM FOXF2 CDNA PCRS. 279
TABLE 6.2: INCOMPLETE FOXF2 CDNA FRAGMENTS 281
FIGURE 6.5: CDNA FRAGMENTS IN RELATION TO THE TRANSCRIPT 284
FIGURE6.6: PEVRF0 EXPRESSION PLASMID 286
TABLE 6.3: SITE DIRECTED MUTAGENESIS THERMAL CYCLING CONDITIONS 289
PRIMERS AND PCR CONDITIONS WHICH PRODUCED THE REQUIRED MUTATIONS WITHIN THE FOXF2 PEVRF0 PLASMID. 289
TABLE 6.4: PRIMERS FOR TESTING THE INTEGRITY OF PEVRF0 CONSTRUCTS 290
FIGURE 6.7: FOXF2 EXPRESSION CONSTRUCT MUTAGENESIS 291
FIGURE 6.8 AMINO ACID SEQUENCE OF FOXF2 AND THE TRUNCATED FOXF2 OF THE PEPTIDE 292
FIGURE 6.9 EVIDENCE OF EMPTY PEVRF0 PLASMID 293
FIGURE 6.10: DIG OLIGONUCLEOTIDE LABELLING EFFICIENCY TESTS 295
FIGURE 6.11: NON SPECIFIC COMPETITOR SELECTION 296
FIGURE 6.12: PROTEIN CONCENTRATION OPTIMISATION 297
FIGURE 6.13 SPECIFIC COMPETITOR ASSAY 298
FIGURE 6.14: SPECIFIC COMPETITOR ASSAY 300
FIGURE 6.15: EMSA OF COS-7 CELL LYSATE FOLLOWING TRANSFECTION WITH FOXF2 CONSTRUCTS 301
FIGURE 6.16: ENLARGED VIEW OF EMSA 303
FIGURE 6.17 OPTICAL DENSITIES OF EMSA BANDS 304
FIGURE 6.18: PTAL-SEAP REPORTER 310
FIGURE 6.19: MULTIMERISATION OF THE SINGLE BINDING SITE OLIGONUCLEOTIDES 312
FIGURE 6.20: FOXF2 REPORTER PLASMID CONSTRUCTION SCHEME 314
FIGURE 6.21: THE DIFFERENT OLIGONUCLEOTIDES USED FOR BINDING SITE INSERTION 315

FIGURE 6.22 TRANSFECTION EFFICIENCY CONTROL DATA 318
FIGURE 6.23 TRANSACTIVATION ASSAY DATA 321
FIGURE 6.24: EGFP WAS USED A A TRANSFECTION CONTROL IN THE TRANSACTIVATION ASSAY 322
FIGURE 6.25: POTENTIAL FOXF2 BINDING SITES ON ENDOGENOUS PTAL-SEAP 324

Table of Contents

Abstract xv
Chapter 1: Introduction 1
1.1 Deletions 1
1.2 Mechanisms of Deletion Formation 2
1.3 Mouse Deletions - Tools for Recessive Phenotype Screens 3
1.4 Alternative Methods of Deletion Induction: Targeted Deletions 7
1.5 The Specific Locus Mutagenesis Test ((SLM) and Mouse Mutagens 9
1.6 Genetic & Physical Mapping 11
1.7 N-Ethyl-N-nitrosourea (ENU) 14
1.8 Phenotyping 16
1.9 Deletion Syndromes of Human Chromosome 6p 18
1.10 The Del(13)Svea36H Deletion 19
1.11 The different Gene Families Within the Del 36 region 21
1.11.1 Vomeronasal receptors 21
1.11.2 Prolactins 21
1.11.3 Butyrophilins 22
1.11.4 Serpins 22
1.11.5 Transcription Factors 23
1. 12 Forkhead Transcription Factors 24
1.12.1 Structure of Forkhead Genes 24
1.12.2 Forkhead Protein - Chromatin Interaction 25
1.12.3 Forkhead Gene Function 27
1.13 The Forkhead Cluster Within the Del36 Region 34
1.14 The Del36H Eye Phenotype 35
1.15 Foxc1 & The Eye 36
1.16 Evidence of at least one other glaucoma locus at 6p25 38
1.17 FOXF2 Associated Glaucoma? 40
1.18 Aims of the Project 43
Chapter 2: Materials and Methods 45
2.1 Materials 45
2.1.1 Antibodies 45
2.1.2 Competent Cells 45
2.1.3 Enzymes 45
2.1.4 Film 46
2.1.5 Kits 46
2.1.6 Membranes 47
2.1.7 Plasmids 47
2.1.8 Radiolabelled nucleotides 48
2.1.9 Stock DNA 48
2.1.10 Precast Gels 48
2.1.11 Solutions 49
2.2 Methods 55
2.2.1 Plasmid Preparation using the Promega Wizard Plus SV Miniprep 55
2.2.2 Preparation of Mouse Genomic DNA 55

Method 1: For DNA extraction from tail tips and whole organs 55
Method 2: For DNA extraction from ear-clips and toe-clips and embryonic tail tips 56
2.2.3 Quantification of DNA Concentration 57
Spectrophotometry 57
2.2.4 Oligonucleotide Primer Design and Synthesis 57
2.2.5 Polymerase Chain Reaction (PCR) 58
Programme 1 Standard PCR 59
Programme 2 Touchdown PCR 60
Programme 3 DHPLC PCR 60
Programme 4 Site Directed Mutagenesis 61
2.2.6 PCR Product purification 61
2.2.7 Agarose Gel Electrophoresis of Nucleic Acids 61
2.2.8 Extraction of DNA from Agarose Gels 62
2.2.9 Random Oligonucleotide Radiolabelling of DNA probes 62
Overgo End-Fill Method 63
Blocking of Repeats in Probe with Mouse Cot-1 64
Hybridisation of Probes 65
Genomic Library Screening by Hybridisation 66
2.2.10 RPCI23 Mouse BAC library. 66
Primary Identification of Genomic Clones by Hybridisation. 66
2.2.11 BAC DNA Vector storage 67
2.2.12 Restriction Enzyme Digestion of DNA 67
2.2.13 Single Stranded Conformational Polymorphism (SSCP) Assays 69
2.2.14 Novel Polymorphic marker Identification 69
DNA Preparation for Novel Marker Identification 69
Novel Microsatellite Identification 70
Novel SNP Identification 71
2.2.15 Production of the Genotyping NED size standard 71
2.2.16 Genotyping 73
2.2.17 Ligation 75
2.2.18 Transformation 75
Transformation using XL-1 blue supercompetent cells. 75
Confirmation of Transformants by Sequencing 76
2.2.19 Sequencing 76
Method 1 77
Method 2 79
Sequencing using the ABI Prism 3130XL Genetic Analyzer 79
2.2.20 Total RNA Extraction from Mouse Eye and Lung Tissues 80
Additional RNA Quantification and quality control 80
2.2.21 First Strand cDNA Preparation 81
2.2.22 Reverse Transcriptase PCR (RT-PCR) 81
2.2.23 Bioinformatics 82
2.2.24 F1 DNA and sperm archive 83
2.2.25 Denaturing High-Performance Liquid Chromatography (DHPLC). 83

2.2.26 Pyrosequencer 84
2.2.27 Mouse Necropsies 84
Eye Extractions for Histological Analysis 84
Lung Extractions for Histological Analysis 85
Head Extractions For Palate Histological Analysis 85
Head Extractions For Eye Analysis 85
Embedding and haematoxylin and eosin Staining 86
2.2.28 Mouse Behavioural Studies 86
Acoustic Startle & Pre-Pulse Inhibition 87
Startle 1 87
PPI 1 88
Startle 2 88
PPI 2 88
2.2.29 Mouse Metabolic Phenotype Studies 88
Metabolites tested with the Olympus AU400 89
2.2.30 Plasmid Preparation 90
Site Directed Mutagenesis 90
Empty pEVRF0 91
Reporter Vector Construction 92
Oligonucleotide Multimerization 92
2.2.31 Probe Dig Labelling 94
2.2.32 Cell Culture Techniques 95
Trypsinisation of Cells 95
Cos7 Cell Passage 96
Cryopreservation of Cells 96
Reviving Frozen Cos-7 Cell Stocks 96
2.2.32a Optimisation of Transfection Conditions 97
2.2.33 Transactivation Assay 98
2.2.34 Reporter Activity Quantification 100
Secreted Alkaline Phosphotase (SEAP) Quantification 100
Luciferase Quantification 100
2.2.35 EMSA 101
Protein Extraction 101
Protein Quantification 102
EMSA Reaction and Analysis 102
1. Bio-Rad Mini-Gel System. Recipes of 5% and 6% were used: 103
2. CriterionTM 5%TBE pre-cast gels (Bio-Rad) 104
2.2.36 Antibody Detection for FOXF2 transfection confirmation 105
Chapter 3: Mapping the Region of Mouse Chromosome 13 Deleted in the Del(13)Svea36H Mutation 107
3.1 Introduction 107
3.2 Physical Mapping of Del36H 109
3.3 Source of Probes used for Hybridisation of the RPCI-23 BAC Library 110
3.4 Construction of the Physical Map 112
3.5 Lethal Recessive Screen 115
3.6 Mouse Strains Used for the Lethal Recessive Screen 116

3.7 Microsatellite Marker Design and Polymorphism Testing 116
False Positive Results 122
3.8 *Satin* and 3H1 Unexpected SSCP Profiles 124
3.9 Marker Redundancy 126
e-PCR 127
3.10 SNP Discovery 128
3.11 Construction of the Genotyping Panels 131
Marker Panel 1 131
Marker Panel 2 134
Optimising product size of genotyping markers 136
3.12 Analysis of Genotyping Data 139
Ambiguous 296bp FAM Product 141
Genotyping 143
Lethality testing 143
Inheritance testing 144
Mapping and Mutant Line Maintenance 145
3.13 Line 91 Recombinant Genotyping 147
3.14 Foxf2 Mutation Detection 149
Chapter 4: Foxf2 Mutation Detection 152
4.1 Introduction 152
Foxf1 and Foxf2 Expression and Sequence Homology 153
Foxf2 in the eye 153
4.2 Why Search for an Allelic Series? 155
The Gene Driven Screen & The ENU Archives 156
4.3 Identification of Mutations with Denaturing HPLC 157
4.4 Optimum Temperature Conditions for Mutation Detection 159
4.5 Amplicon Descriptions and dHPLC Runs 162
Mutations 168
4.6 Mutant Peak Analysis 171
Mutation 1 - GSK 1B11 171
Splice Site Mutation - MRC 31H8 172
V-F mutation - MRC 18C1 172
W174R mutant peak - GSK: 14H3 172
4.7 Mutation Discovery Discussion 178
G-T mutation (Silent 1) at position 231 GSK archive 1B11 178
Splice Site Mutation. MRC archive 31H8 179
V-F mutation. MRC18c1 182
Silent 2 Mutation (MRC 25H1) 183
W174R Mutation (GSK 14H3) 184
Foxf2_1_2a with W174R and Silent 1 185
Mutation Detection Probabilities 189
The Mutations 191
Chapter 5: Phenotyping of Mouse Mutant Foxf2^{W174R} 193
5.1 Introduction 193
5.2 Heterozygote Screening 197
5.3 Appearance 198

5.4 Behaviour 199
5.4.1 Behavioural Observations and Response to Simple Stimuli 199
5.4.2 Light Dark Box Test 205
5.4.3 Spontaneous Alternation 208
5.4.4 Swim Test 208
5.4.5 Neuromuscular test (Grip Strength test) 208
5.5 Startle Tests 209
5.5.1 Problems with Programmes Startle-1 and PPI-1 210
5.5.2 Startle-2 211
5.5.3 Analysis of the Startle Data 213
5.6 Pre–pulse Inhibition (PPI) 220
5.6.1 PPI-2: Retesting of the Initial Male Cohort 224
5.6.2 Is there a difference in the PPI Response of Foxf2^{W174R} Mutants? . 225
5.7 Metabolism Tests 225
5.7.1 Blood Tests Using the Olympus AU400 Clinical Chemistry Analyser 225
5.7.2 Urine Dipstick Tests 230
Ascorbic Acid 233
Protein 234
5.8 Intra-Peritoneal Glucose Tolerance Tests 235
5.9 Weight Screen 238
5.9 Weight Screen 238
5.10 Eye Analysis 240
Development of the Ocular Anterior Segment 240
Eye Screen of heterozygotes aged 330-340(days) 252
5.10.2 Eye Histology of Homozygote Embryos 252
5.11 Foxf2^{W174R} Homozygote Analysis 254
5.11.1 Genotype Ratios 254
5.12 Pathology/Necropsy Reports 259
5.12.1 Heterozygote Pathology Screen 259
5.12.2 Screen of Homozygotes at day 3 259
5.13 Cleft Palate Tests 260
5.14 Lung Histology 262
5.15 Prognosis 263
Chapter 6: Molecular Phenotyping of Two FOXF2 mutations 266
6.1 Introduction 266
6.1.1 FOXF2 Nuclear Localisation & Mode of Action 270
6.1.2 The Foxf2 Binding Sequence 272
6.2 EMSA 273
6.2.1 Mouse cDNA Expression Vector Construction 276
6.2.2 A Human FOXF2 Expression Construct 285
6.2.3 Vector Mutagenesis 288
6.2. 4 Mutant Plasmid Production 288
Primer Sets Used for Plasmid Integrity Testing 290
6.2.5 Production of the Empty pEVRF0 Expression Plasmid 293
6.3 Transactivation Assay 309

6.3.1 Construction of the Reporter Plasmid 309
6.3.2 Control Levels in EMSA Transfections 318
6.3.3 SEAP Data 319
6.4 Potential pTAL-SEAP Endogenous FOXF2 Binding Sites 323
6.5 Conclusions 326
6.5.1 Potential Sources of Error that could Explain the Data 327
6.5.2 Future Work 329
Chapter 7: Discussion 331
7.1 Physiological Phenotyping 333
7.2 Characterisation of the Del36H Eye Phenotype 335
7.3 Alternative Candidates for the Lethal Recessive Screen Mutations 337
7.4 Heart Defects 340
7.5 Metabolic Phenotype 341
7.6 Startle and Foxf2 343
7.7 Molecular Phenotyping 344
7.8 Future Work 346
References 349
Appendices for Chapter 3 367
Appendix 3.1 367
Appendix 3.2 374
Appendix 3.3 377
Appendix3.4 379
Appendix 3.5 382
Appendix 3.6 401
Appendix 3.7 402
Appendix 3.8 403
Appendix 3.9 404
Appendix 3.10 405
Appendix 3.11 407
Appendix 3.12 408
Appendix 3.13 409
Appendices for Chapter 5 410
Appendix 5.1 410
Appendix 5.2 411
Appendix 5.3 412
Appendix 5.4 413
Appendix 5.5 415
Appendix 5.6 416
Appendix 5.7 416
Appendix 5.8 417
Appendix 5.9 417
Appendix 5.10 419
Appendix 5.11 420
Appendices for Chapter 6 428
Appendix 6.1 428
Appendix 6.2 428

Appendix 6.3 ..... **429**
Appendix 6.4 ..... **430**
Appendix 6.5 ..... **431**
Appendix 6.6 ..... **431**
Appendix 6.7 ..... **432**
Appendix 6.8 ..... **433**
Appendix 6.9 ..... **434**
Appendix 6.10 ..... **435**

Mapping and Phenotyping Genes in the Del(13)Svea36H Region of Mouse Chromosome 13

Abstract

Del(13)Svea36H (also known as Del36H) is a 12.7Mb gene-rich region that has been deleted in mouse chromosome 13. It contains numerous disease loci as well as showing conserved synteny to two regions of human chromosome 6p which can be missing in some deletion syndromes. The mouse deletion causes homozygous embryonic lethality but is hemizygously viable. It has therefore been used as a tool to screen for lethal recessive mutations that could help to identify developmental genes that are responsible for the symptoms of the associated human diseases

A Physical map of the region was constructed in order to facilitate the mapping of mutations arising from the lethal screen. Eleven lethal lines have been mapped to varying degrees of accuracy, including one lethal mutation that exhibited circulation defects which localised to Sox4.

The forkhead transcription factor *Foxf2* was considered a potential candidate for embryonic lethal mutations within the interval. To complement the lethal recessive screen, an attempt was made to identify an allelic series of mutations in the gene. ENU mutagenised mouse archives were screened for novel *Foxf2* mutations. Three potential functional mutations were discovered - two coding mutations ($Foxf2^{W174R}$ and $Foxf2^{V412F}$), and a splice site mutation.

Foxf2^{W174R} heterozygotes exhibited thinning of the iris stroma, alongside hypoplasia of the trabecular meshwork and canal of Schlemm and a reduction in the irido-corneal angle. Homozygote eyes at E18.5 show no signs of the developing ciliary body which were seen in wildtype littermates.

Homozygote carriers survive for several days, significantly longer than the 18hr survival time observed in the knockout, suggesting a hypomorphic mutation. In contrast to *Foxf2* knockouts, homozygotes for the W174R mutation do not possess a cleft palate or malformed tongue.

Mapping and Phenotyping Genes in the Del(13)Svea36H Region of Mouse Chromosome 13

Chapter 1: Introduction

1.1 Deletions

Large chromosomal deletions can occur in the human population and commonly result in mental retardation, morphological malformations or restricted growth when the aberration is inherited [1-3], or cancer in the case of deletions in somatic cells [4]. Deletions are one of several forms of gross chromosomal anomaly, others include trisomy, inversion, duplications and translocation mutations [5]. An estimated 20-50% of all human conceptions carry such chromosomal abnormalities, although a much lower 0.6% of live births carry them [6].

Macrodeletions have traditionally been regarded as those deletions which can be visualised cytogenetically by light microscope with giemsa staining, whereas microdeletions were those deletions which could not be detected with this method. This distinction however, has since become less relevant with the advent of techniques such as FISH (Fluorescent In-Situ Hybridisation) and more recently with the higher throughput array-CGH (Comparative Genomic Hybridisation) [7]. The more appropriate term is perhaps "Contiguous Deletion Syndrome" which was coined by Ballabio [7] to describe deletions which spanned at least 2 genes and produced a defined clinical phenotype. This group of diseases could be considered a subset of the "Contiguous Gene Syndromes" that Schmickel conceptualised [8] or

"segmental aneusomy syndromes" [9] which also include the inversions, duplications and translocations of fragments containing several genes.

The phenotypes which are observed in carriers of these deletions are likely to be caused by the reduction in gene dosage of those genes within the region for which dosage is critical for normal function [10].

Large chromosomal deletions which are not associated with a detectable clinical phenotype have been detected in the human population [11]. However, it is possible that effects may have gone undetected due to the absence of a relevant environmental challenge.

1.2 Mechanisms of Deletion Formation

A lot can be learned about the mechanisms of deletion and other chromosomal rearrangements by the comparative analysis of closely related species such as chimpanzee and man [12-15]. There are two mechanisms which are thought to cause the majority of known deletions. These are: – 1) Non Allelic Homologous Recombination (NAHR) and 2)Non Homologous End Joining (NHEJ) [16]. Recombination driven by either of these mechanisms can result in the formation of a deletion when two different alleles on the same chromosome recombine during meiosis with the resultant loss of the region between the two alleles.

NAHR is the predominant mechanism for deletion and duplication syndromes [16, 17] whereby alleles with high levels of sequence identity, usually Low Copy Repeats (LCRs), are the substrate for erroneous recombination[18, 19]. Several other chromosomal features have also been implicated in the process however, and in particular other repeat elements such as α-satellites, transposons and chi-sites [16]

Those deletions which are thought to be caused by non homologous end joining (NHEJ) tend to have widely scattered breakpoints compared to those formed by the tightly clustered rearrangements seen in NAHR [16]. NHEJ is well documented in tumourogenesis but its involvement in hereditary deletions is not as apparent as is NAHR.

1.3 Mouse Deletions - Tools for Recessive Phenotype Screens

Mouse deletion mutants can model human deletion syndromes. There has been significant rearrangement during the evolution of both organisms, so large deleted regions are rarely contiguously syntenic, but large regions of synteny are present.

Deletions can be induced at various sites across the mouse genome, without destroying the viability of the hemizygous deletion carriers [20]. These carriers often demonstrate a range of developmental problems but are able to carry the deletion and survive and breed (albeit often with a reduced viability) [20, 21]. These mice represent an excellent opportunity for regional screens to identify novel recessive

mutations, and thus the potential for a greater understanding of the functionality of genes within these regions.

X-ray induced deletions and ENU have been effectively combined in a powerful mutation discovery approach which was first used on mice by Eugene Rinchik[22]. He was inspired by the success of similar methods that employed ethyl methanesulfonate (EMS) to identify novel genes in *drosophila* [23].

This breeding strategy utilised the female F1 progeny of ENU mutagenised BALB/cR1 males that were homozygous for the recessive albino mutation (c/c). These were crossed to a male carrier of an X-ray induced deletion on chromosome 7 [22]. Hemizygous Lethal mutations were identified by a birth ratio of fewer than 1:30 albino coated offspring from individual females. Rinchik suggested that this approach could be applied to a region in order to obtain saturation mutagenesis i.e.: at least one mutation in every gene in the deleted region.

In their review of phenotype based mutagenesis Screens, Schimenti and Bucan highlight several advantages of the regional deletion based recessive screen over the whole-genome alternative [24]:

1. The whole genome recessive screens require 3 mating stages instead of the two stages that are used in deletion screens for recessive mutations.
2. In the third stage of the whole genome screen, only 12.5% of the progeny are likely to be homozygous for any given mutation, in contrast to the 25% of

progeny likely to be hemizygote mutation carriers following the final stage of deletion screens. This is because only 50% of the parental mice in the final mating stage will be mutation carriers during the whole-genome method (see figure 1.1).

3. The backcrosses required for the whole genome approach also means that if several mutations are present, the results of phenotype analysis are likely to be confused.
4. The use of visible markers within deleted regions means that lines carrying lethal recessive mutations can be easily identified by the absence of the visible trait. The whole-genome method is unlikely to detect these highly informative mutations.
5. This technique also allows the detection of recessive alleles which are effectively already "pre-mapped" to a deleted region.

Although a large chromosomal region is "targeted" in deletion screens, this method does not have the inherent bias of gene-targeted approaches because no assumptions are made about which genes within the region are likely to disrupt normal function.

In addition to their use as a tool for the identification of recessive mutations, mouse deletions can also model chromosome anomalies in humans [25] . They have the potential to increase the depth of knowledge about the genes responsible for the human phenotype and how the genes' displacement effects the genetic interaction between genes within the region. Deletions also provide a unique opportunity to

analyse the cumulative effect from altered expression levels of more than one gene within a region when closely linked genes are involved in the same biological process [26].

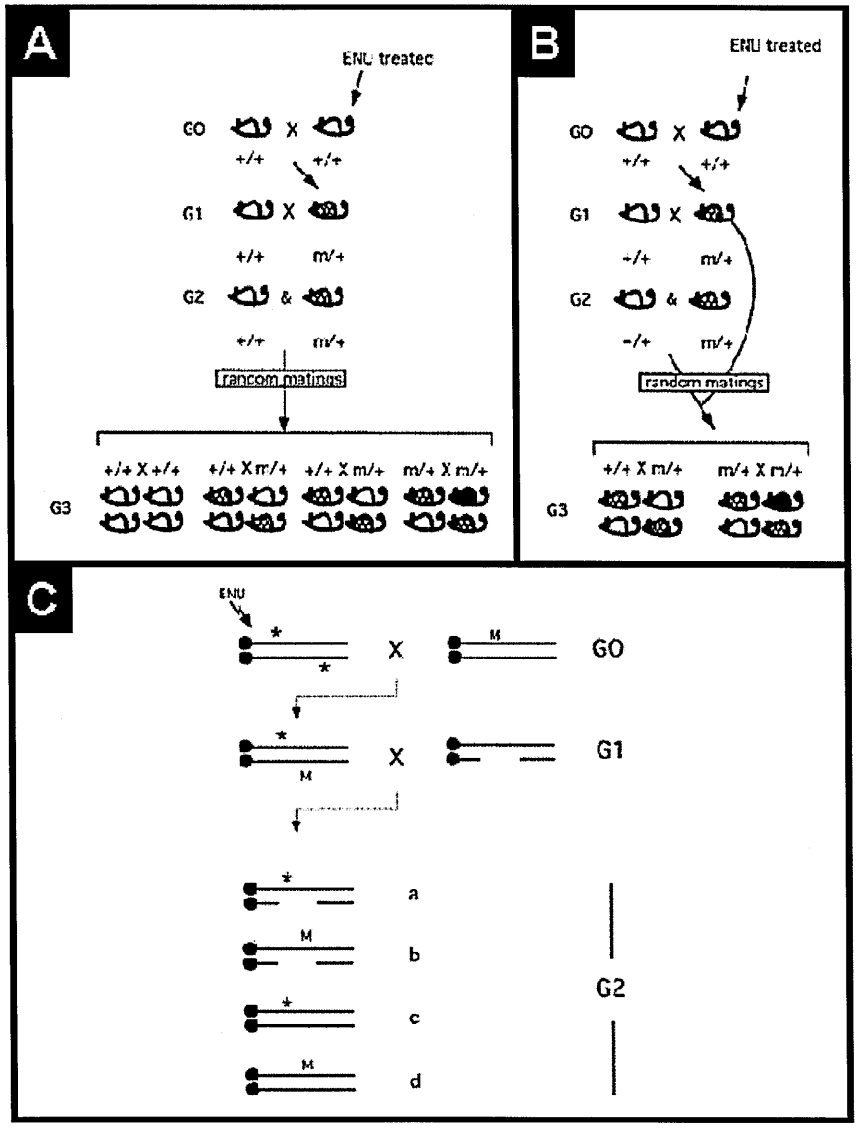


Figure 1.1: Whole Genome & Deletion Mutagenesis Screens
Schemes representing the whole genome and Deletion based mutagenesis recessive screens. Both approaches involve the initial exposure to G0 mice with ENU which induces mutations (represented by "m" in A & B, and by "*" in C). The G1 offspring are then mated to wildtypes to generate G2 litters which are either intercrossed with siblings (A) or backcrossed to their fathers (B). Mutagenised individuals in the regional screen (C) are crossed with animals that contain a dominant visible marker (M) to produce mutation (*) carriers that are subsequently crossed with carriers of a large chromosomal aberration such as an inversion or deletion. The absence of the visible marker in the progeny from this cross suggests the possible presence of a lethal recessive mutation. Diagram adapted from Schimenti and Buchan [24].

1.4 Alternative Methods of Deletion Induction: Targeted Deletions

One shared characteristic of ENU and X-ray mutagenesis is that both systems cause random mutations within the genome; the experimenter has no control over the region affected. An alternative technique which was pioneered by Ramirez-Solis & colleagues, has utilised a Cre-lox system to induce specific mutations within any region of interest in the genome of mouse ES cells.

Ramirez et al described an approach for creating deletions whereby loxP sites were inserted into the flanking regions of an intended deletion site in embryonic stem cells. Cre recombinase induces recombination between these sites [27]. Different types of cis- chromosomal rearrangement (i.e. deletion, inversion, duplication) are dependent on loxP orientation. These recognition sequences contain an 8bp asymmetric sequence at their centre which confers directionality. Deletions are likely to form when the two loxP recognition sites are aligned in the same direction and on the same chromosome. This is because when two such sites align, a loop forms between the two sites and this is subsequently removed from the chromosome on recombination (figure 6.2).

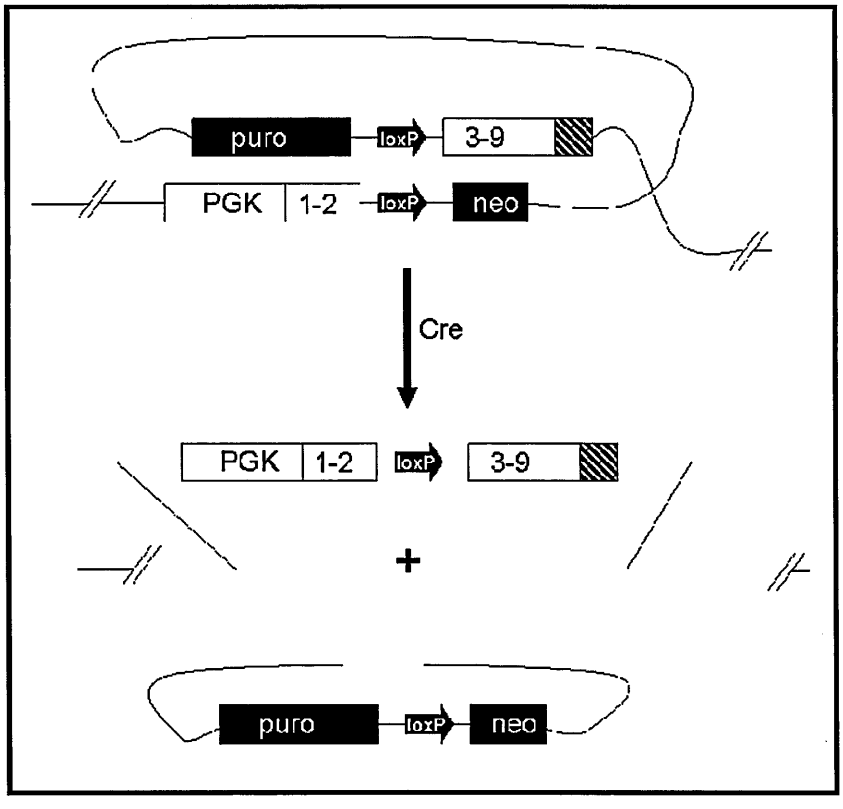


Figure 1.2: Cre-lox system for Mutation Induction
The Cre-lox system can be used to form deletions provided that the loxP sites are oriented in the same direction and are on the same chromosome. Cre recombinase induces recombination between the loxP sites. The features labelled "1-2" and "3-9" represent exons of Hprt (Hypoxanthine phosphoribosyl transferase). This is a gene which allows the breakdown of hypoxanthine and therefore allows survival in HAT media. It is only viable when the exons are correctly positioned in relation to each other, which only occurs following recombination between the lox sites, they therefore provide a system of positive selection for recombination. The unidirectional nature of the LoxP sites means that their orientation can affect the type of chromosomal rearrangement which they induce and other alignments would instead result in inversions, duplications or translocations. PGK is a promoter element, puro and neo are resistance genes that are required for gene targeting of the Hprt cassettes. This image was adapted from Ramirez-Solis et al [27]

The Cre-lox system has also been used to induce targeted point mutations [28], so can also produce the full range of different allelic forms. This approach though, requires the screening of a large number of embryos and there is an inherent bias in any targeted approach because of the assumptions that must be made about which regions, genes, domains and nucleotides are important enough to target.

The Trans-allelic targeted meiotic recombination (TAMERE) strategy has since been used to modify this method. Mice which carry different loxP loci can be crossed with each other to facilitate either deletions or other chromosomal rearrangements [29]. The availability of these mice would reduce the number of electroporation and selection stages for the insertion of novel loxP sites in individual crelox chromosome rearrangement experiments [29].

1.5 The Specific Locus Mutagenesis Test ((SLM) and Mouse Mutagens

William Russell designed the specific locus mutagenesis (SLM) test to determine the heritable gene mutation rate caused by radiation in germ cells [30]. However, this technique had much further implications and formed the basis of experiments which led to the discovery of several hundred "specific locus mutations" at Oak Ridge National Laboratory in the USA and at the Medical Research Council at Harwell [31]. These mutations later formed a crucial framework for the genetic mapping of the mouse.

This test utilised a mouse (The T-stock mouse) which harboured seven recessive mutations that produced distinctive visible traits. These loci were: nonagouti (a) on chromosome 2, the brown (b) allele of Tyrp1 on chromosome 4, the chinchilla (C^{ch}) allele of Tyr on chromosome 7, the dilute (d) allele of Myo5a on chromosome 9, pink-eyed (p) on chromosome 7, the piebald-spotting (s) allele of EdnrB on chromosome 14 and the short-ear (se) allele of BMP5 on chromosome 9.

Wildtype males were X-ray irradiated before crossing with T-stock females. Mutations which disrupted the wildtype allele at any of the seven specific loci resulted in a failure to complement the recessive allele so that the visible traits were expressed thereby identifying mutation carriers.

Russell noted that the mutations discovered with this approach frequently mutagenised both the d and se loci and he suggested that these individuals may carry large mutations which spanned the region between the loci [30]. This theory was given greater credence by the construction of a more detailed linkage map between d and se which charted the relative positions of 8-9 independent functional units [32] and it has since become clear that many of the X-ray induced mutations arising from the SLM strategy were large chromosomal inversions, deletions or translocations which affected multiple genes[33] [34].

1.6 Genetic & Physical Mapping

The Mouse genome sequencing project was initiated in 1996 to complement the human sequence. It provided an opportunity to improve previous mouse mapping data. Sequence would facilitate physical positioning of all previously identified loci and allow the precise determination of physical distances between them.

By 1996, a dense mouse genetic map of 6580 polymorphic SSLP markers had been developed [35]. The known mouse genome rapidly expanded with the use of sequence tagged sites (STS), radiation-hybrid panels[36], and YAC libraries[37]

As human genome sequencing began in the early nineties, YACs (Yeast Artificial Chromosomes) made up the most comprehensive of the human and mouse physical maps [38]. By 1995 various genome-wide STS-based human maps were completed, including a genetic map, a YAC map, and an integrated YAC-RH genetic map. These maps provided a framework for positioning additional loci [39]. However YACs have several associated disadvantages. Deletions and rearrangements within YAC sequences are frequent, they are prone to a high rate of chimaerism, and DNA extraction from the yeast host can be problematic [40].

Because of these problems, alternative systems were identified for the cloning, mapping and sequencing of the genome. For example, P1 bacteriophage derived

plasmids: PACs [41], and BACs (Bacterial Artificial Chromosomes)- a vector system which utilises the innate F-factor conjugation plasmid of E. coli [42].

Although PACs are much more stable than YACs, they only have a maximum insert size of 100kb[42]. One disadvantage of BACs is that there is no positive selection for recombinants, meaning that identification of recombinant clones is required by colony hybridisation[41]. On balance though, when the hierarchical approach has been used for large scale physical mapping, BACs have been considered the best option due to both their large insert size and relative stability[43].

There were two strategies for sequencing the human genome: The whole genome shotgun approach that was championed by Craig Venter of Celera, breaks the entire genome into sequenceable fragments[44]. For a time, the publicly funded sequencing centres were in direct competition with Celera but had already decided upon a method known as hierarchical shotgun sequencing, in which the genome is first broken into overlapping large-insert clones[45]

The whole genome shotgunning approach is the simpler of the two, and gave rapid early coverage, but genomes with highly repetitive sequences can be problematic[43]. Hierarchical shotgun overcomes this problem because it is a localised assembly method - the number of repeat copies in each assembly is reduced and it is possible to compare large regions of overlap between the clones. A

combination of both approaches was decided upon for sequencing of the mouse genome[43].

It is vital that sequence-ready maps should contain contigs (several clones that contain overlapping sequence) which are of sufficient depth that the clones can be validated by several other overlapping clones [40]. This helps to reduce the chance of including sequence from clones which have undergone rearrangement which would incorporate incorrect sequence data into the construct.

1.7 N-Ethyl-N-nitrosourea (ENU)

Another major contribution to mouse genetics by William Russell, in addition to the SLM test, was the discovery of the potent mutagenic effect of Ethylnitrosourea (ENU) in mice [46].

Russell had observed that in contrast to SLM tests with X-rays, tests with ENU did not produce mutations which occurred at both the se and d loci simultaneously – he therefore suggested that ENU may cause mutations within individual genes rather than large multi locus regions [46]. Stronger evidence for this arose when the peptide sequence of haemoglobin α-chain 5 in an ENU mutagenised DBA/2 mouse was shown to carry a single amino acid substitution (His-Leu) which suggested an A-T point transversion in the gene [47].

ENU is an alkylating agent which forms ethyl free radicals that can bond to nucleobases and cause point mutations in DNA[48, 49]. It preferentially reacts with AT base pairs [50] and since the discovery of the first ENU induced point mutation, the majority of reported mutations in mouse have been AT-TA or AT-GC base pair transitions [48, 51]

These point mutations do not necessarily knockout the total functionality of a gene as is often the case with mutations caused by other mutagens or gene traps[52]. Instead, a range of different effects can be observed in genes mutagenised at a single base. Hypomorphs, which result from reduced gene dosage or protein activity are the

most common ENU induced allele [51]. These are followed by the null alleles that other less specific mutagens tend to cause. ENU can also result in hypermorphic mutations that indicate an increase of gene dosage or protein activity; neomorphs where a protein has acquired a novel function; and antimorphic alleles which are dominant negative effects that can be caused by competitive inhibition of normal protein function [51] [53] [48].

This range of different allelic forms that can be caused by point mutations can potentially provide vastly more information about the activity of a gene and its product than a knockout (null allele) alone could provide[49]. Null alleles certainly can give a very useful window of insight into the role of a gene when its not functionally redundant, but a detailed analysis of protein function and of interaction with other factors is not possible without other allelic forms.

The potential for ENU was highlighted to the scientific community with the discovery of the clock gene thanks to the presence of an ENU induced circadian rhythm mutation on mouse chromosome 5 [54]. The excitement surrounding this insight into such a complex and what had been up to then, little understood molecular system provided the impetus for an explosion of ENU mutagenesis experiments during the 1990s and into the new millennium.

Two large scale genome wide ENU mutagenesis screens for dominant mutations were independently initiated by Steve Brown at Harwell [55] and Rudi Balling at the

Institute of Experimental Genetics in Neuherberg, Germany [56]. A systematic phenotyping approach was taken; following general observations of each animal screened, a wide range of tests were performed to optimise the potential detection of any morphological, biochemical or behavioural mutants.

1.8 Phenotyping

Mouse mutations are of limited use without a detailed characterisation of the phenotype of the carrier. The ultimate aim of mouse mutagenesis is the linkage of genetic factors with physical or behavioural traits, and for this to be realised accurate phenotyping is essential. SHIRPA is a comprehensive and standardised series of tests designed to highlight a wide range of behavioural, physiological, and metabolic mouse phenotypes [57]. It has been used both for both phenotype screens in the high-throughput analysis of large numbers of individuals which have been subjected to random mutagenesis [58], and genotype screens when the phenotype of mice with a mutation in a known gene is interrogated [59].

There are many factors which could potentially obscure a real association between a mutation and a mouse abnormality. The common inbred strains are known to vary extensively in various phenotype tests, a phenomenon which is being explored with the mouse phenome project [60, 61](http://www.jax.org/phenome). The behavioural traits which are detected by SHIRPA, often show significant differences between the common inbred strains, including learning and memory [62], locomotor activity [63], and anxiety [64]. This suggests that a mixed background could mask potential

behavioural phenotypes. Repeated backcrossing to a selected strain can minimise these effects, but even following multiple backcrosses of over 10 generations, large regions that are undisturbed by recombination can flank the gene of interest, and frequently contain in excess of 300 genes [65].

In addition to these genetic considerations, a wide range of environmental factors can drastically affect results, particularly in behavioural screens. Noise, light, diet [66], cage environment [67], husbandry techniques, experimenter [68] and pathogenic environment are all factors which are likely to vary between mouse facilities.

Differences in experimental protocols can also have an important influence on the results obtained. The Eumorphia project (http://www.eumorphia.org) was initiated to try to standardise working practices so that results between laboratories can be reliably and directly compared. A hierarchical phenotyping platform is under ongoing development which consists of a set of standard operating procedures (SOPs) that are intended as a reference point for robust phenotyping protocols for the mouse community [69]. Following validation studies, many of the tests are now consistent between centres, but others still require the elimination of variables which prevent reliable data comparisons [70].

1.9 Deletion Syndromes of Human Chromosome 6p

Deletions of sections of chromosome 6p in humans are rare. In 2004, Mirza et al counted 43 cases in the medical literature [2], since then at least 11 new cases have been reported [2, 71-77]. It therefore seems that these syndromes may have previously been under-reported, but with improved cytogenetic techniques the frequency of small deletion detection is increasing [78].

Davies and colleagues were the first to note that there are two distinct groups of 6p deletion syndrome: 6p25 deletions, which is the 6p region at which deletion frequencies are highest, and variable interstitial deletions between 6p22-24 [2, 79, 80].

Clinical findings from the 6p25 region, the most frequently deleted region of the 6p syndromes, include corneal opacity, malformation of various structures of the ocular anterior chamber, hypertelorism (widely spaced eyes), ptosis (drooping eyelids), optic nerve abnormalities, hydrocephalus, Dandy-Walker malformation, heart defects, deafness and mental retardation [2, 73].

The symptoms of the interstitial syndromes have included orofacial clefting, short neck, clinodactyly (curved fingers), syndactyly (fusion of fingers), and brain, heart and kidney defects [2, 79].

Although these deletions can provide a wealth of evidence relating to the function of genes in the deleted regions, direct comparisons of phenotype with deletion span can be misleading because of other factors affecting phenotype such as the environment of the womb during gestation, parental origin and mutations in other genes with related function. This has been noted as a particular source of difficulty in the characterisation of the 6p25 associated eye phenotypes which have highly variable penetrance levels [2].

There has been a concerted effort to characterise the correlation between map positions of various 6p25 deletions and the associated eye phenotype. This interest has been primarily fuelled by the phenotypic link to the well characterised Foxc1 mouse mutant, congenital hydrocephalus (ch) and the Foxc1 knockout [81]. These mutants die with hydrocephalus, skeletal abnormalities and Axenfeld Rieger anomaly which involves malformations of the anterior segment of the eye that can lead to glaucoma.

1.10 The Del(13)Svea36H Deletion

Del(13)Svea36H is a deleted region of mouse chromosome 13 [82] with conserved synteny to human chromosome 6p21.3– 6p23 and 6p25[21]. The deletion can be detected cytogenetically and was initially estimated to result in the removal of approximately 12% of the chromosome at G-bands 13A2-13A3.2 [82]. Subsequent cytogenetic analysis however, which reviewed the deleted interval in 9 animals, suggested an expanded interval at G-bands 13A3-13A5 and was 20.1% of the

chromosome [21]. This is quoted as equating to 25.6Mb with the most accurate estimation of chromosome 13 size at that time (corrected to 24.18Mb with the more reliable chromosome 13 size taken from the Ensembl v50 sequence assembly).

The later estimate was given greater credence by high resolution genetic mapping with the EUCIB (European Collaborative Interspecific Backcross) resource [83] that indicated a similar size of 12.5cM, equating to 26.5Mb (corrected to 25.03Mb with Ensembl v50). However, following the precise localisation of the end-points of the deleted region (J. Raggoussis, personal communication) and attainment of the completed sequence for this region, a more reliable size estimate of 12.7 Mb has been made. This figure is closer to the initial cytogenetic size estimate of ~12% of the chromosome which would have equated to ~15.28Mb with the earlier whole-chromosome size estimate (or corrected to ~14.43Mb with chromosome size data from Ensembl v50).

There is a high density of genes in the Del36H region – with 201 known genes [84]. This has led to speculation that the Del36H region may have undergone high frequencies of recombination during its evolution, which would have encouraged gene duplication [84]. Regional Features that increased recombination rates during evolution are likely to affect size estimations determined by genetic mapping, because inter-loci distances are calculated by recombination frequencies with these methods.

1.11 The different Gene Families Within the Del 36 region

A detailed annotation of the Del36 region was carried out using manual annotation and comparative sequence analysis [84]. The high gene density of the region is largely due to the expansion of three gene families: The prolactins, the serpins and the vomeronasal receptors. There has also been an expansion of a butyrophilin cluster on human 6p, in which seven butyrophilin genes in the human region are represented in the syntenic del36H region by only two butyrophilin mouse genes. Several histone gene clusters are present which can also be found on the syntenic human 6p region.

1.11.1 Vomeronasal receptors

This family of chemo-sensory proteins are thought to mediate the response of the vomeronasal organ to chemical stimuli [85]. In humans the vomeronasal organ regresses during foetal development, although this is not the case in mice [86]. This is reflected in the wide disparity in the size of this gene family between these two species and there has been a significant vomeronasal expansion in Del36H when compared with the human syntenic regions [84].

1.11.2 Prolactins

The prolactin family of genes have undergone species specific gene duplication at an extraordinary scale – there is only one member of the family in human chromosome 6p, but 26 active genes in the syntenic region of the mouse Del36H region, including

three pseudogenes [84]. Wiemers et al predicted that there were 26 fully functional genes [87] but Mallon pointed out that three of four of the genes predicted by Wiemers on the basis of cDNA and EST data had not been mapped and may be the result of sequence or copy-number polymorphism in the region which may have appeared as individual genes due to the use of multiple mouse strains [84]. Prolactins are expressed in the uterus, placenta and anterior pituitary. They are responsible for the regulation of maternal-foetal interactions during pregnancy and the induction of milk production after birth [87].

1.11.3 Butyrophilins

Butyrophilins are members of the immunoglobulin superfamily. They are associated with the membranes of lipid droplets in milk and are involved in the regulation of droplet secretion. The malformed lipid droplets in Btn1a1 mouse knockouts suggest that they are also important for the integrity of the membrane structure of these droplets [88]. 50% of the pups that suckled from Btn1a1(-/-) mothers died within 20 days of their birth.

1.11.4 Serpins

The term serpin is an abbreviation of "serine protease inhibitor". Many serpins are active protease inhibitors, but the term has been applied on a structural rather than a functional basis, and some of them do not demonstrate any enzyme inhibiting activity [89]. All of the serpins in the human syntenic region to del36H however, are from the ovalbumin (ov)-like serpin clade, and are either cysteine or serine protease

inhibitors [89, 90]. Ov-serpins are frequently involved in tumour development, inflammation and apoptosis control. Many of the ov-serpins that are present in the Del36 region have localised expression in the placenta, although there are also a number with more widespread expression patterns in other, varied organs and two with apparently ubiquitous expression in adults [90].

1.11.5 Transcription Factors

Transcription factors are a large and diverse group of genes that code for proteins which either repress or facilitate the activation of DNA transcription. They have modular structures and the large majority contain a DNA binding domain. Some of the basal transcription factors do not bind to DNA but contribute to the pre-initiation complex that interacts with RNA polymerase II. Another common feature of transcription factors are trans-activating domains which contain binding sites for attachment to other proteins such as transcription co-regulators.

1. 12 Forkhead Transcription Factors

Forkhead box (Fox) transcription factors were named after the spiked ectopic head structures that formed in association with the diminished guts of mutagenised *Drosophila* embryos [91]. A domain within the causative forkhead (fkh) gene was identified that had high levels of sequence homology to Hepatocyte Nuclear Factor 3 factor (HNF-3), suggesting that they may be members of a larger family of transcription factors [92].

The known forkhead family, members of which all contain this highly conserved DNA binding domain, has greatly expanded since the discovery of these first members and there are currently 42 known Fox genes in humans, 16 in drosophila and at least 4 in yeast [93]. The presence of the family in yeast is evidence which supports the hypothesis that fungi and animals form an evolutionary supergroup - the opisthokonts, and are linked by a more recent ancestry than are other eukaryotes [94]. The apparent correlation between the complexity of an organism and the number of Fox genes it carries led Carlson to suggest that increasing complexity of body plan may drive the expansion of the Fox gene family [94].

1.12.1 Structure of Forkhead Genes

The first forkhead protein analysed to be structurally characterised by X-ray crystallography was rat HNF-3 (now known as FoxA3) [95]. The phrase "winged helix" was used to describe the helix-turn-helix structure of the forkhead DNA

binding domain. This consists of three clustered helices on an antiparallel beta-sheet structure which forms the base of two extended loops or "wings". Subsequent NMR analysis of other forkhead proteins suggested an additional alpha helix in these molecules[96, 97] but it was later realised that the predicted presence or absence of these structures was probably due to differences in interpretation rather than sequence variation [98]

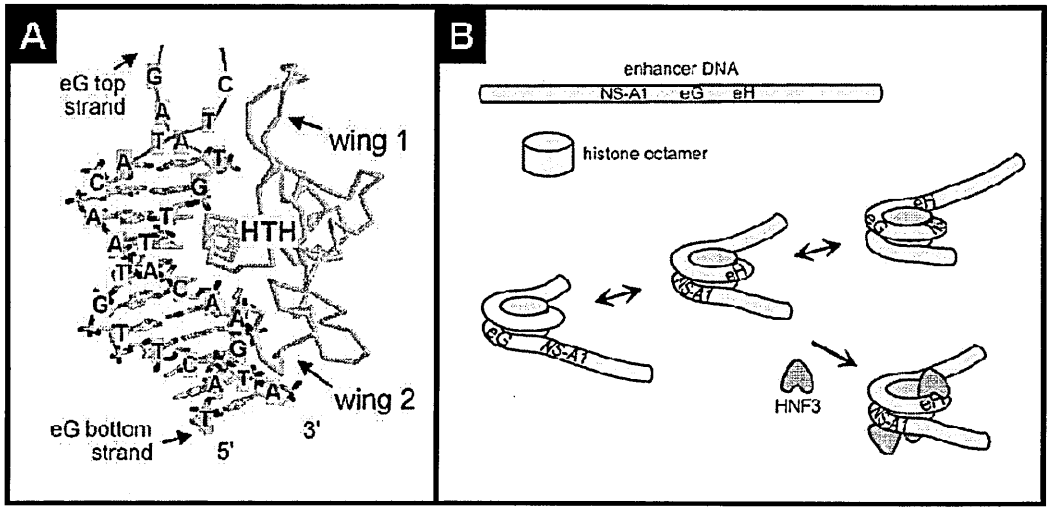


Figure 1.3: Forkhead proteins binding to DNA
Image A = structural prediction of how the structural features of the Foxa3 protein may fit into the helical form of its DNA binding sequence. This image was taken from Cirillo et al, 2007 [99]. Image B = The envisaged alignment of HNF3 (now known as Foxa3) on DNA when the DNA is wrapped around histones as chromatin. This image was taken from Cirillo et al, 1999[100].

1.12.2 Forkhead Protein - Chromatin Interaction

Cirillo and colleagues found that unlike many other transcription factors, the FoxA proteins bind with greater affinity to DNA wrapped around histones as chromatin than they do to free DNA [100]. Recent data suggests that the wing-like structures of the protein are important for keeping the protein on one side of the DNA helix, allowing space on the other side for the histones or other transcription factors [99].

A domain at the C terminal end of the Fox A proteins is thought to bind to histones 3 and 4 within the nucleosome. Cirillo believes that this binding combined with the strong affinity of the forkhead domain to the specific DNA site, is probably the beginning of a process which sees the opening of the chromatin structure to allow access to the DNA [101].

Much of the work investigating the chromatin – forkhead interaction has been carried out with the structurally well characterised FoxA proteins, but Carlsson and Mahlapuu suggested that a common theme uniting forkhead genes may be the mechanisms by which they interact with chromatin [94]. This could well be the case as the DNA binding domain is highly conserved throughout the forkhead family. However FoxA – footprinting experiments showed that the C terminal domain was important for binding to histones 3 and 4 of the nucleosome to facilitate the chromatin opening and therefore access to the DNA. Other transcription factors that require access to chromatin would also probably need histone binding regions, but sequence outside the forkhead domain is highly variable between forkhead genes and the C terminal region studied was thought to be unique [101]. The exact method of chromatin entry is therefore likely to vary somewhat between the different members of the forkhead family.

1.12.3 Forkhead Gene Function

Despite the highly conserved forkhead domain, forkhead transcription factors are an extraordinarily diverse group of genes with key roles in the development, homeostasis, stress response and cell cycle control of a wide range of tissue types. Many of these genes also have a broad pattern of expression individually and mutations within them often cause highly pleiotropic effects.

Some forkhead transcription factors can act as key modulators of the cell cycle. Foxm1 is responsible for the upregulation of CDC25A, a phosphatase which activates the transition from phase G1 into S phase [102] and without Foxm1, mitosis does not occur. It also down-regulates $p21^{CIP1}$ and $p27^{kip}$, which are molecules that function as inhibitors of the cell cycle shift [102].

$p21^{CIP1}$ is regulated by Foxk1 in myogenic stem cells [103] [104] and by Foxg1 in cortical progenitors [105]. The FoxO clade also functions to regulate cell cycle arrest by the regulation of $p27^{kip}$ [106, 107].

Fox genes in the developing Embryo

Foxa2 is a key component of the pathways that form the node and the notochord. Without it, severe defects in gastrulation, neural tube patterning and gut morphogenesis cause early embryonic lethality [108]. Foxa2 also contributes to the correct functioning of the foregut endoderm, in combination with Foxa1. Many other

tissues of endodermal origin such as liver, pacreas, lungs and prostate also require Foxa1 and 2 for normal development. Glucose metabolism can also be affected in FoxA mutants due to the abnormal functioning of the liver, pancreas and adipose tissue [109].

Fox genes are essential for correct skeletal formation. Foxc1 deficiency causes skull and thoracic defects [110]. In addition, mesenchymal cells fail to differentiate into cartilage, even with the addition of BMP2 or TGFB1 [81]. Foxc1 is expressed in the paraxial mesoderm at early gastrulation, and then subsequently in the presomitic mesoderm and the first branchial arch [111].

Fox gene expression is very tightly localised during the formation of the branchial arches. Foxc2, Foxd1, Foxd2, Foxf1 and Foxf2 expression is lost from a range of facial primordia when hedgehog responsiveness is incapacitated in neural crest cells [112]. These forkhead genes have also been reported to be induced by Shh in somites and foregut [113-116].

Initiation of forkhead gene expression by hedgehog transcription factors is a common theme in the literature and it is possible that the forkhead susceptibility to the effects of hedgehog proteins may be a common feature of the entire gene family.

Forkhead gene induction by Sonic hedgehog is necessary for aortic arch development, and is also required for expression of Foxa2 and Foxc2 in the

pharyngeal endoderm and head mesenchyme [117]. Foxa2 expression in the notochord and the floorplate of the neural tube is maintained by a positive feedback loop with sonic hedgehog. Sonic hedgehog also interacts with Foxa1 and Foxa2 in the branching morphogenesis of the mammalian lung [118].

Fox genes in the heart

Deficiency of Foxp1 can result in Atrial Septal defect, defective myocardial proliferation and abnormal outflow tract septation [119]. The forkhead genes Foxc1 and Foxc2 are also involved in the regulation of cardiac outflow tract septation in association with the known cardiac regulator Tbx [117, 120]. Foxp1, Foxc1, and Foxc2 are expressed in the cardiac neural crest [117, 119-121]. Foxn4 regulates tbx5 which controls the development of the AV boundary [122]

Lung

Foxp1 and Foxp2 repress the mouse CC10 promoter and, to a lesser extent, the human surfactant protein C promoter – thus implicating them in the regulation of epithelial lung gene transcription[121]

Glucose Tolerance

The FoxO and FoxA clades play fundamental roles in the appropriate formation of pancreatic islet cells and liver. They are also involved in the regulation of many metabolic genes such as are present in the hepatic gluconeogenesis pathway – which is important in the prevention of hypoglycemia during fasting [123].

Foxo proteins regulate the entry of cells into the G1 cell cycle and are critical for cell survival and proliferation. Foxo1, Foxo3a, and Foxo4 contain three phosphorylation sites for insulin activated protein kinase pathways [123].

Even closely related forkhead genes can have apparently completely divergent functions. There is mounting evidence that Foxp2 is critical for the correct development of speech and language, following the discovery of a family in which a single amino acid change (R553H) co-segregated with severe verbal communication difficulties in otherwise intellectually normal individuals [124]. The closely related Foxp3 though functions as a lineage specification factor in regulatory T cells, and is important in the maintenance of self-tolerance [125]. There does not appear to be any overlap in the function of these two genes despite a whole protein sequence homology of 54% and a forkhead domain homology of 75% (Ensembl v50).

The variety of tissues and processes in which forkhead genes are involved (See table 1.1) would suggest that the expansion of this family was more indicative of a useful general transcriptional mechanism than of any property that makes them useful for the development, maintenance or activity of specific tissues.

The eye though, is the system in which most Fox related human disorders have been identified to date [126]. It is a delicately balanced system and seems to be prone to suffering effects from alterations in Fox gene dosage [10, 127-129].

The anterior segment of the eye, which is formed by a complex process of interactions between ectodermal, neuroectodermal, and neural crest cells appears at particular risk of such alterations. Disturbances in the balance of this process can result in glaucoma, cataract, leucoma, or other visual defects.

Forkhead mutations in human disease were first identified in association with FOXC1 (previously known as FKHL7) in patients with Axenfeld-Rieger syndrome and glaucoma hypoplasia [130, 131].

This discovery was followed by the discovery of a number of other mutations in various forkhead genes that also resulted in eye disorders. Mouse Foxe3 is expressed in the lens epithelium following a wider distribution in the developing lens. The dysgenetic lens (*dyl*) allele of *Foxe3 prevents DNA binding*. The phenotype is equivalent to Peters' anomaly in humans and includes central corneal opacity, lens and iris hypoplasia, keratolenticular adhesion, absence of secondary lens fibres and severe cataracts. Foxe3 is essential for the proliferation of the lens epithelium and also acts to prevent apoptosis and inappropriate differentiation of its precursor cells. In addition to defects of the anterior segment that are reminiscent of those seen in FOXC1 mutations, some FOXC2 mutations cause abnormal eye-lash growth that can result in damage to the cornea [132]. FOXL2 mutations cause eye lid abnormalities in occasional association with ovarian disorders [133]

Gene	Vertebrate Mutant/Disease Phenotypes
FoxA1	Growth retardation, hypoglycaemia & reduced glucagon production in mouse null mutants [134].
FoxA2	Mouse null - absent node & notochord; foregut development severely impaired [108]. Hypoglycaemia with conditional null Foxa2 mutations in the pancreas [135].
FoxA3	Mouse null - inefficient glucose transport from hepatic cells and hypoglycaemia during fasting [136]
FoxB1	Mouse null - growth retardation. Failure to nurture pups. CNS Defects [137]
FoxC1	Mouse null - hydrocephalus, skeletal, ocular, renal & cardiovascular defects [81, 110, 138] . Human missense, frameshift, deletion and protein truncation mutations - anterior eye segment defects, glaucoma & Axenfeld-Rieger anomaly [131, 139].
FoxC2	Mouse null - skeletal, cardiovascular, lymphatic & ocular defects in mice [120, 140]. Human missense & frameshift mutations - lymphedema & distichiasis syndrome, ptosis, ocular anterior segment anomalies & cleft palate [132, 138].
FoxD1	Mouse null - Renal failure[141]
FoxD2	Mouse null - Partially penetrant renal anomalies [142].
FoxD3,	Mouse null - loss of epiblast cells, expansion of proximal extraembryonic tissues, and an ectopic anterior visceral endoderm [143, 144].
FoxD4,	Human point mutation - Cardiomyopathy, obsessive-compulsive disorder, and suicidality co-segregated with a FOXD4 mutation in a single 3 generation family [145]
FoxD5 (FoxD4L1),	Xenopus null- Immature neural cell fate; Inhibition of neural differentiation [146].
FoxD6 (FoxD4L3),	N
FoxE1	Mouse null - Cleft palate & Aberrant thyroid development in mice [147]. Human missense - as in mouse abnormalities alongside Choanal atresia [148].
FoxE2,	Described by Ensembl as a synonym of FoxE1
FoxE3	Mouse null - fusion of the lens, cornea and iris. Cataract formation [149, 150]. Human frameshift mutation - Eye anterior segment dysgenesis & cataracts [151].
FoxF1	Mouse null - absence of vasculature in the allantois or yolk sack. Also lung and foregut malformations and defects in cell adhesion [116, 152].
FoxF2	Mouse null - cleft palate and malformed tongue [153]; megacolon, colorectal muscle hypoplasia & agangliosis [154]. Human insertion mutation - glaucoma [155]
FoxG1	Mouse null - cerebral hemispheres of the brain are smaller than wildtypes [156].
FoxH1	Mouse null - the node, notochord, prechordal mesoderm and definitive endoderm all fail to develop [157, 158].
FoxI1	Mouse null - Inner ear malformations: Deafness & balance problems [159].
FoxJ1	Mouse null - Ciliogenesis defects and left-right asymmetry in mice [160].
FoxJ2,	Human - Foxj2 is upregulated during heart failure [161].
FoxJ3,	Mouse null - reduced post-natal growth. Decreased fat and type I muscle fibres but increase in type IIa/b muscle fibres. Inefficient muscle regeneration following injury. Increased myoblast proliferation rates. [162].

Gene	Vertebrate Mutant/Disease Phenotypes (Continued)
FoxK1	Mouse null - restricted growth. Defect in the proliferation and differentiation of myogenic stem cells [163].
FoxK2,	Human Embryonic Kidney (HEK-293), mouse fibroblast (NIH3T3) and mouse pituitary (NtT20) cells - apoptosis in proliferating cells. Overexpression prevents apoptosis [164].
FoxL1	Mouse null - structural abnormalities of the stomach & intestine. Uncontrolled epithelial cell proliferation [165]. Increased propensity to tumour formation [166].
FoxL2	Human missense, frameshift mutations & protein truncation & extension mutations: Blepharophimosis, Ptosis, Epicanthus, Ovarian failure (BPES syndrome). [133, 167] [167]
FoxM1	Mouse null - Polyploidy in cardiomyocytes and hepatocytes [168]. Cardiovascular defects [169]. Failure to enter mitosis [170]. Overexpression of mouse Foxm1 accelerates the development, proliferation and growth of prostate tumours [171].
FoxN1	Mouse null - thymus fails to develop, hairless [172]. Human protein truncation mutation - T cell deficiency. Alopecia. Malformed nails [173, 174].
FoxN2	N
FoxN3	Human oral squamous cell carcinoma cells - reduced FOXN3 expression [175].
FoxN4,	Mouse null - early postnatal lethality or noticeable body size reduction. Fewer amacrine and horizontal cells in the retina [176].
FoxN5 (FoxR1),	N
FoxN6 (FoxR2),	N
FoxO1 (FoxO1A),	Mouse null - defective angiogenesis [177]. Mouse Foxo1 Haploinsufficiency restricts the probability of insulin resistance. Overexpression of mutant Foxo1 in mouse liver and pancreas can cause diabetes [178]. Human chromosomal deletion of FOXO1 - common in prostate cancer.
FoxO3 (FoxO3A, FOXO2, FOXO6),	Mouse null - Abnormal ovary development. Early female infertility [179].
FoxO4	No obvious phenotype in knockout mice [179]or known human diseases.
FoxP1,	Mouse null - severe defect in early B cell development [180]. Foxp1+/- mouse heterozygous mutation compounded with Foxp2 null: Severity of Foxp2-/- lung phenotype is worsened. Oesophageal muscle development is disrupted [181].
FoxP2	Human point mutations & chromosomal translocations - speech and language disorder [124]. Mouse null - Defects in lung alveolarisation. Lung defect severity is increased by the additional presence of a single Foxp1 null allele (Foxp2-/-;Foxp1+/-). Oesophageal muscle development is also disrupted in Foxp2-/-;Foxp1+/- embryos [181].
FoxP3	Mouse frameshift mutation. - overproliferation of CD4+CD8- T lymphocytes [182]. Human deletion mutations cause immune dysregulation, polyendocrinopathy and enteropathy (IPEX syndrome) [183].
FoxP4,	N
FoxQ1	Mouse intragenic deletion mutation - defect in hair differentiation [184].

Table 1.1: Phenotypes of mutations in vertebrate forkhead-box genes.
List of Phenotypes associated with vertebrate forkhead genes. Phenotypes listed are primarily sourced from induced mouse mutations and human disease symptoms.
N = No data available. Genes in brackets = Fox named synonyms.

1.13 The Forkhead Cluster Within the Del36 Region

There are 8 clusters of forkhead genes that have been identified within the human genome [126]. One such cluster contains the genes FOXC1, FOXF2 and FOXQ1 and is located at 6p25 in humans and the del36H region of chromosome 13 in mouse. A paralogous cluster is located at human 16q24 (mouse chromosome 8) containing FOXC2 and FOXF1 in place of their close paralogues FOXC1 and FOXF2. There is no direct FOXQ1 paralogue but a forkhead gene from a different clade, FOXL1, is present.

Mazet and co-workers have carried out comparative analysis of these genes in different bilaterian groups. FoxF, FoxC, FoxQ and FoxL are all represented singly in amphioxus on one tightly linked cluster. Mazet et al were therefore able to construct a potential model for the evolution of this cluster before the divergence of the bilaterians [185]. They suggest that the replication of two closely linked genes and maintenance of their relative positions suggests that the paralogues are the result of a block duplication event in early vertebrates. The loss of one FoxQ gene at one locus and one FoxL gene at the other, perhaps caused by their redundancy, could explain the formation that can now be seen in mouse and human.

Expression studies by Mazet showed that during the development of the endo-mesoderm these genes were expressed sequentially [186], and the order of expression is matched by the physical positions of the genes on the chromosome, a pattern reminiscent to that of hox clusters in somite differentiation [187]. Mazet

makes the convincing argument that this pattern in endomesoderm formation could be a selective force for the maintenance of the cluster, and may explain its persistence in most bilateral organisms.

Many of the phenotypes associated with deletions of chromosome 6p25 have been attributed to FOXC1 because many of the phenotypes associated with the deletion syndrome were detected as equivalent phenotypes in Foxc1 knockout mice [81].

1.14 The Del36H Eye Phenotype

The Eyes of Del36H mice were analysed following the initial observation that they had a "gummy" appearance [82]. Ruth Arkell and co-workers later noted that the eyes were open at birth and 40% of carriers had irregularly shaped pupils in one or both eyes. Older mice had opaque eyes which were severely scarred and frequently infected. The "gummy" appearance of eyes was described as secondary to the eyes being open at birth [21].

Mutations in several forkhead genes cause a wide spectrum of eye disorders [126]. A forkhead cluster is present in the Del36H region that includes the gene Foxc1 which has a strong association with defects of the ocular anterior chamber [81]. Foxf2 – a neighbour of Foxc1 and also a member of the cluster, is expressed in the eye from ~day 13 of mouse embryonic development [188]. These are therefore immediate candidates for the analysis of the eye defect in the Del36H mutants.

Another potential factor that may contribute to the Del36H eye phenotype is that prolactins, of which there are a large number in the deleted region [84], have been shown to inhibit angiogenesis in the retinas of rats [189]. It is possible that overproliferation of blood vessel cells could result in malformation of the eye(s).

1.15 Foxc1 & The Eye

The mouse Foxc1 null homozygote mutation causes death at birth and the mice suffer from hydrocephalus and a range of skeletal abnormalities. In the anterior segment of the eye, a range of morphological anomalies were present including iris hypoplasia, malformed eyelids and incorrectly aligned corneal cells [81].

Mice that were heterozygous for the mutation possessed a disorganised and undifferentiated trabecular meshwork, irregular pupils, iris hypoplasia, small or absent Schlemms canal, and a displaced Schwalbes line [138].

In vivo cell fate mapping has since demonstrated that neural crest cells are present from day E10 in the optic cup and lens vesicle of healthy wildtype mice [190]. Coordinated interactions between the neural ectoderm and periocular mesenchyme are necessary for the formation of the highly specialised structures of the ocular anterior chamber. TGFβ2 activates the expression of *Foxc1* and *Pitx2* which are critical for the differentiation and proliferation of neural crest cells.

The neural crest cells eventually develop to form the iris stroma, ciliary body, trabecular meshwork, choroid and primary vitreous of the mature eye [190]. It therefore seems likely that the Foxc1 mutations are affecting the developing eye by interference with the correct differentiation, proliferation or activity of these cells.

The abnormal structures in the anterior segment of Foxc1 null mutants possess insufficient levels of extracellular matrix (ECM). The ECM that was present was not the characteristically organised beam structure that is typical of this component of the anterior segment, but was instead disorganised and with reduced levels of collagen and elastic tissues. However despite these abnormalities – which affect the eye drainage system, most of the mice did not have increased intraocular pressure.

ECM of structures in other systems such as the sternocostal cartilages, sternum and kidneys are also disrupted in Foxc1 mutants, and Smith suggested that this may indicate that Foxc1 has a role in the regulation of the synthesis or organisation of ECM. He suggested that the similar phenotype seen in Foxc2 individuals could also indicate a role in this pathway.

Collagen is a key constituent of the ECM. Col18a1 knockout mice, which cannot produce type 18 collagen, have an abnormal ciliary body and iris [191]. Type 8 collagen is also important in eye development and when the two subunits are disrupted, there is a decreased proliferation of endothelial cells which leads to a thinning of the cornea and Descemet's membrane [192]. These studies indicate that

collagen levels are important for the correct development of the anterior segment. Collagen should perhaps therefore be considered a likely downstream target of transcription factors that are associated with eye malformations.

Douglas Gould and colleagues have predicted a set of potential genetic pathways of iridocorneal angle development which centres on the conversion of tyrosine to dihydroxyphenylalanine (L-Dopa) [193]. Gould et al believe that this molecule is likely to be important in the process because it is a precursor to catecholamines which are developmental regulators [194]; and because it can be linked to known pathways for many of the genes involved in angle formation such as PAX6, LMX1B, PITX2 and 3 and for Bmp4, Tgfb2 which demonstrate similar phenotypes to Foxc1 and Foxc2. Foxc1 can also be linked to this pathway since the recent finding that PITX2 binds to and inhibits the activity of the FOXC1 protein [127].

1.16 Evidence of at least one other glaucoma locus at 6p25

There have been several previous reports of additional glaucoma loci within the 6p25 region. A paper published shortly after glaucoma was first linked to Foxc1 [130], mentioned four glaucoma patients whose condition was linked to 6p25 but in whom mutations were absent from the coding region of Foxc1 [131]. In two of these cases, Foxc1 was also excluded by familial recombination mapping.

Gould highlighted another patient with an unbalanced translocation between 6p25 and 4p14 [80], in which the breakpoint was between two markers that surrounded FOXF2 but was distal of FOXC1. This raised the possibility that FOXF2 may have been monosomic and FOXC1 disomic in this individual. However the heterozygous marker (D6S967), which lies between the two genes, is significantly closer to FOXF2 than was the hemizygous marker (D6S942) (see table 1.2).

Locus	Ensembl.v50 Position		Zygosity
	Start	End	
D6S942	765305	765544	Hemizygous
Foxf2	1335068	1340831	-
D6S967	1428847	1428985	Heterozygous
Foxc1	1555680	1559126	-

Table 1.2: Markers used by Gould et al to Characterise a 6p Translocation
Updated Positions of markers that were used by Gould [80] to characterise a 6p unbalanced translocation which had a breakpoint distal to Foxc1 that resulted in loss of the 6p terminus.

In addition to these findings, Davies reported the case of a patient with a 6p interstitial deletion that was at least 1200kb proximal of Foxc1 [195], and contained neither the FOXC1 or FOXF2 genes. This data suggests that if Foxf2 is a glaucoma locus there may also be a third locus in the 6p25 region, although as Kleinjan suggests, the expression of Foxc1 may have been altered in this case by the disruption of a regulatory element [196].

1.17 FOXF2 Associated Glaucoma?

Cheryl Gregory-Evans, Helena Vieira and colleagues identified a poly-glycine expansion in the FOXF2 gene of a family of glaucoma sufferers [155]. A 3bp insertion mutation (Gly302-306ins3bp) in the second sub-region of the AD2 transactivation domain of FOXF2, has been implicated in a glaucoma which was present in six generations of a family [155] who develop symptoms in their 5^{th} to 6^{th} decade (C. Gregory Evans, personal communication). Glaucoma is an eye disorder which is characterised by damage to the optic nerve and retinal nerve fibres as a result of an increased intraocular pressure.

Pilocarpines are commonly used as a topical treatment for glaucoma [197], they can be used to stimulate the ciliary muscle to increase fluid drainage thereby reducing pressure in the eye, thus preventing optic nerve damage. However, individuals that harbour the FOXF2 mutation do not respond to pilocarpine treatment (C. Gregory Evans, personal communication) - implying that there may be a ciliary muscle abnormality.

The family also suffer from iris hypoplasia which involves thinning and patchiness of the iris stroma. (see figure 1.4). In combination, these symptoms would suggest generic problems with the development of the anterior segment of the eye.

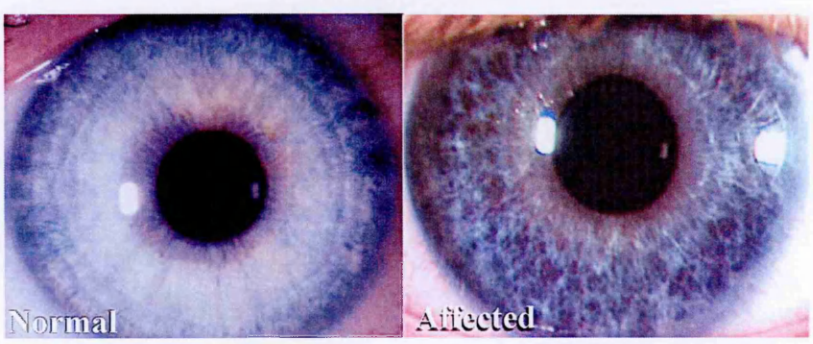


Figure 1.4: Transillumination Defect of the human eye
(supplied by C. Gregory-Evans): A healthy human eye has an outwardly visible thick, white iris stroma, whereas the stroma of the affected patient is thinned and patchy.

Foxf2 heterozygous knockout mice exhibited anterior chamber anomalies that shared many properties with human irido-goniodysgenesis [155]. These included a narrowing of the irido-corneal angle and defects in the trabecular meshwork and canal of Schlemm. This is a phenotype resulting from the aberrant migration of the neural crest cells that form the anterior segment of the eye [190].

Foxf2 was shown to be expressed in the neural crest cells in this region (figure 1.5). Each structure which is involved with fluid drainage from the eye is formed by these neural crest cells and therefore any defect in their migration is likely to impact upon the drainage of aqueous humour and thus also on intraocular pressure.

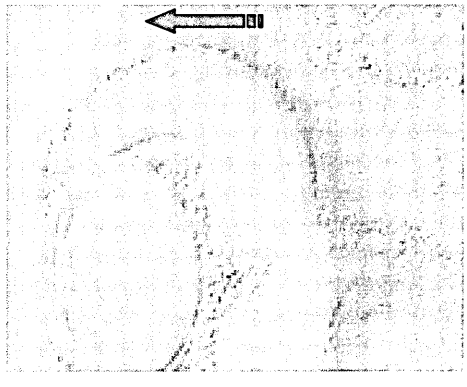


Figure 1.5: Periocular mesenchyme in a mouse eye at E12.5
An in-situ of Foxf2 in mouse eye development at E12.5. Purple staining is located at the level of the periocular mesenchyme (that contains migrating neural crest cells). Cells migrate forward in the direction of the arrow. (in-situ hybridisations performed by, and picture supplied by H. Vieira & C. Gregory-Evans)

Genes that contribute to ocular anterior segment dysgenesis (ASD)[193] often either effect the activity or migration of neural crest cells or they have roles in the developing lens. Lens genes are thought to contribute to ASD because the lens controls many of the processes required during the formation of the anterior segment[198, 199]

1.18 Aims of the Project

The initial aim of this project was to facilitate the sequencing of the Del36H region by fine mapping, and to contribute towards the characterisation of the region's genes. The most immediate benefit of constructing a fine map of the Del36H region was the accuracy with which it could subsequently be used for positional cloning of mutations from the lethal recessive screen – an experiment for the functional characterisation of the genes in the region.

Regional phenotype screens are designed to uncover novel mutations within known chromosomal rearrangement intervals, eliminating the requirement for a preliminary mapping stage following confirmation of the existence of a mutation. However, Del36H is large at 12.7Mb and gene rich (~200 genes), so mutations that are discovered will require further localisation before the sequencing of candidate genes can begin. In order to facilitate this process, the region will need to be searched for novel polymorphic markers. These are then to be utilised to genotype recombinant individuals arising from the screen.

One of the most striking overlaps between the phenotype of the Del36H mouse and patients presenting with 6p deletion syndromes are the Axenfeld-Rieger anomalies and the equivalent mouse eye defects. FOXC1 is a known locus for Axenfeld-Rieger anomaly in humans and anterior eye defects in mice [81, 200] but various studies reported another potential locus in 6p25 [80, 131]. The Del36H mouse presents an ideal opportunity to study this phenotype in greater detail.

The potential involvement of another gene that is physically close and closely related to FOXC1 has recently been highlighted: H. Vieira and C. Gregory-Evans discovered a FOXF2 polyglycine expansion in a family of glaucoma sufferers [155]. This has provided the incentive to undergo a search for novel mutant alleles of Foxf2 in the Harwell ENU mouse mutagenesis archives [201, 202]. Mouse mutant carriers of a Foxf2 mutation with a glaucomatous phenotype would provide strong supporting evidence to further implicate FOXF2 as a key genetic factor in heritable glaucoma.

A wide range of organs and tissues are affected by the 6p deletion syndromes, and Forkhead genes are often highly pleiotropic. There is therefore an incentive to thoroughly examine as many potential phenotypes of Foxf2 mutants as possible. The use of a broad phenotyping platform for the analysis of physiological, behavioural and metabolic abnormalities is necessary to achieve this goal.

Chapter 2: Materials and Methods

2.1 Materials

2.1.1 Antibodies

Anti-Digoxigenin-AP Fab fragments (Roche)

Anti-Rabbit IgG (whole molecule)-Peroxidase (Sigma)

FOXF2 Antibody 1: Rabbit Anti-Human FOXF2 Polyclonal Antibody. 5mg/ml. (Abcam®)

FOXF2 Antibody 2: Mouse Anti-Human FOXF2 Polyclonal Antibody, Unconjugated serum (Abnova® Corporation)

2.1.2 Competent Cells

XL1-Blue Supercompetent Cells (Stratagene)

JM109 competent cells (Stratagene)

2.1.3 Enzymes

Amplitaq Gold™ - 2.5u/µl - (Applied Biosystems)

BigDye ® v3.1 (Applied Biosystems)

DNA Polymerase I, Large (Klenow) Fragment – 5000u/ml - (New England Biolabs)

Klenow – 1u/µl – (Amersham Life Sciences)

Pfu Turbo® DNA Polymerase – 2.5u/µl - (Stratagene)

Positive Control Placental Alkaline Phosphatase (BD Biosciences)

0.1mg/ml

Thermoprime Plus DNA Polymerase 2.5u/µl - (ABgene)

Restriction enzymes. All restriction enzymes were purchased from New England Biolabs:

Sfc I, Bsr I, Dde I, SnaB I, Mlu I, Bgl II, Hind III, Kpn I, Dpn I, Sal I, Xba I, BamHI

Sensiscript^{TM} (Qiagen)

Superscript^{TM} (GibcoBRL)

T4 DNA Ligase - 400u/ml (New England Biolabs)

2.1.4 Film

Hyperfilm ECL Chemiluminescence Film (Amersham Biosciences)

Scientific Imaging Film (Kodak)

2.1.5 Kits

DC Protein Assay kit (Bio-Rad)

DIG DNA Labelling Kit (Roche)

DIG Wash and Block Buffer Set (Roche)

DNeasy® 96 Tissue Kit (Qiagen)

ECL^{TM} **Western Blotting Analysis System** (GE Healthcare)

EndoFree® Plasmid Maxi kit (Qiagen)

Great EscApe^{TM} **SEAP Detection Kit** (BD Biosciences)

MightyLight™ Rluc Assay Kit (Novagen, Merck)

Multiplex PCR Kit (Qiagen)

Nucbuster Protein Extraction kit (Novagen, Merck)

Proteoextract™ Subcellular Proteome Extraction kit (Calbiochem)

Qiaquick® Gel extraction kit (Qiagen)

Qiaquick® PCR purification kit (Qiagen)

Wizard® Plus SV Miniprep (Promega)

2.1.6 Membranes

Hybond N+ (GE Healthcare)

0.45 μm nitrocellulose/Filter Paper Sandwich (Invitrogen)

2.1.7 Plasmids

pEGFP – C1 (Clontech)

pEVRF0 & FOXF2(Gly302-306ins3bp) insert. A mutant FOXF2 (Gly302-306ins3bp) expression construct was supplied by Cheryl Gregory Evans. This was initially constructed by Marika Hellqvist and co-workers [203] who inserted a truncated FOXF2 cDNA sequence (missing 39 amino acids from the 5' end) into a pEVRF0 plasmid [204]. The Gly302-306 insertion was induced with site directed mutagenesis by H.Vieira [155].

pGEM-T Easy (supplied with the pGEM-T Easy Cloning System, Promega)

pIRES2-EGFP (Clontech)

pMLuc-3 CMV Positive Control (Merck – Novagen)

pTAL-SEAP (Clontech).

2.1.8 Radiolabelled nucleotides

[α-^{32}P]-dCTP (Amersham)

2.1.9 Stock DNA

1kb DNA size marker ladder (GibcoBRL)

100ng/ µl. In ration of 1:4 OrangeG:dH_2O

Lambda DNA at 500µg/ml (New England Biolabs)

Mouse CotI DNA - 1mg/ml - (Invitrogen)

Salmon Sperm DNA (10mg/ml)

2.1.10 Precast Gels

Criterion™ Precast Gel. 5%TBE. (Biorad)

NativePAGE™ Precast Gel. 4-16% Bis-Tris (Invitrogen)

NuPAGE™ Precast Gel. 4-16% Bis-Tris (Invitrogen)

2.1.11 Solutions

Laboratory grade reagents and distilled water (dH_2O) were used to make up solutions. Autoclaving of these solutions, where stated, was carried out at a temperature of 120°C and a pressure of 38 PSI for 20 minutes.

Acetonitrile (Fisher)

Ampicillin Stock

50mg/ml ampicillin (sodium salt, Sigma). -20°C

Betaine (Sigma)

5M stock

BigDye Reaction buffer (5X) (Applied Biosystems)

400mM Tris-HCl; pH9.0, 10mM $MgCl_2$

BSA (100X) (New England Biolabs)

10mg/ml

Buffer IV (ABgene)

Chloramphenicol Stock

34mg/ml chloramphenicol dissolved in EtOH. -20°C.

Church-Gilbert Hybridisation Solution

1mM EDTA, 500mM Na_2HPO_4 (pH 7.2), 7% (w/v) SDS, 1% (w/v) BSA

Cryopreservation Media

10%DMSO in Foetal Bovine Serum

CSPD Chemiluminescent substrate (Roche)

Deoxyribonucleotide Triphosphates (10X)

2.5mM dATP, dCTP, dGTP, dTTP (Boehringer Mannheim and Biogene). Stored at 4°C

DEPC H_2O

0.1% (v/v) Diethyl pyrocarbonate (DEPC) in dH_2O

1ml DEPC in 1l ddH_2O. Parafilm on lid and incubated 1hr @ 37°C – then autoclaved.

Dimethyl sulphoxide (Sigma)

EMSA Binding buffer (4X)

400 mM KCl, 80 mM HEPES, 2 mM DTT, 0.8 mM EDTA, 80% glycerol, pH 8.

HotstarTaq reaction buffer (10X) (Qiagen)-15mM $MgCl_2$, KCl, Tris-HCl, $(NH_4)_2SO_4$; pH8.7

Dulbecco's minimal essential medium (DMEM)

1X DMEM + Glutamax with 4.5g/l glucose, without pyruvate (Gibco).

Dulbecco's Phosphate Buffered Saline (DPBS)

1XDPBS (Gibco)

EDTA (Ethylenediaminetetra-acetic acid)

0.5M EDTA at pH8.0. Autoclaved.

Ethidium bromide

10mg/ml. 4°C in dark.

Foetal Bovine Serum (Gibco)

Glycerol (80%)

Genejuice® (Novagen)

Growth Media for Cell Culture

10% Foetal Bovine Serum (FBS) & 1% Penicillin-Streptomycin solution in DMEM

Hepes buffer (5X)

100mM Hepes (Sigma)

HepG2 (Abcam®)

Human hepatocellular liver carcinoma cell line whole cell lysate

Hi-Di™ Formamide (Applied Biosystems)

Isopropylthiogalactosidase (IPTG)

24mg/ml. Filter sterilized. 4°C

Karnovsky fixative (50%)

2.5% glutaraldehyde, 2% paraformaldehyde, 0.05M sodium cacodylate (pH 7.4)

Loading buffer 6X (Merck – Novagen)

Loading buffer for Gel based sequencing

5:1 deionised formamide: 25mM EDTA with 50mg/ml blue dextran. -20°C

Loading buffer for SSCP Gel Electrophoresis

95% (v/v) formamide, 10mM EDTA (pH7.5), 1%bromophenol blue, 1% (w/v) xylene cyanole.

Luria's Broth (LB) and LB agar mix

10g/l Bacto-tryptone, 5g/l Bacto yeast extract, 5g/l Sodium chloride (Difco Laboratories Ltd)

For LB agar mix: 1.5% agar was added (Difco Laboratories Ltd). 4°C

Magnesium Chloride

$MgCl_2$ (Applied Biosystems)

25mM at -20°C

Neutral Buffered Formaldehyde 10% (Surgipath Europe Ltd)

Formaldehyde 3.8 – 4.0% (w/v) [205]

NZY+ broth

10g/l NZ amine (casein hydrolysate), 5g/l yeast extract, 5g/l NaCl. pH 7.5.

Autoclaved. Filter sterilised additives: 12.5ml of 1M $MgCl_2$, 12.5ml of 1M $MgSO_4$, and 20ml of 20% (w/v) glucose.

Orange G: Loading buffer for agarose electrophoresis.

0.25mg/ml Orange G (sigma) in 30% (v/v) glycerol

PCR Reaction buffer II (Applied Biosystems)

Penicillin – Streptomycin solution -10,000u/ml penicillin & 10mg/ml streptomycin in 0.9% NaCl (Sigma)

Pfu DNA polymerase reaction buffer (Stratagene)

Phospate Buffered Saline (PBS)

0.8% (w/v) NaCl, 0.02% (w/v) KCl, 0.115%(w/v) Na_2HPO_4, 0.02% (w/v) KH_2PO_4. Autoclaved

Phospate Buffered Saline – Tween (PBS-T)

0.8% (w/v) NaCl, 0.02% (w/v) KCl, 0.115%(w/v) Na_2HPO_4, 0.02% (w/v) KH_2PO_4. Autoclaved. 0.1% Tween.

Paraformaldehyde (Sigma)

(4%) in PBS or DPBS

Proteinase K (Sigma)

10mg/ml in 10mM $CaCl_2$. -20°C

Q Solution (Qiagen)

Restore™ Western blot stripping buffer (Pierce)

RT Buffer (Invitrogen)

200 mM Tris-HCl, pH 8.4, 500 mM KCl

Sodium Acetate

3M NaOAc (pH5.2. adjusted with glacial acetic acid)

Sodium Dodecyl Sulphate (SDS) – (stock)

10% (w/v) in H_2O

Standard Saline Citrate (SSC) – (20Xstock)

3M NaCl, 0.3M Na_3citrate.$2H_2O$ (pH7.0 adjusted with 1M HCl)

Tail Buffer

500mM Tris Buffer, 100mM EDTA, 100mM NaCl, , 1% (w/v) SDS

Tris-acetate EDTA Electrophoresis buffer (TAE)

40mM Tris-acetate, 1mM disodium EDTA (pH8.0)

Tris-borate EDTA Electrophoresis buffer (TBE)

89mM Tris, 89mM boric acid, 2mM EDTA

Tris Buffered Saline – Tween (TBS)

0.8% (w/v) NaCl , 0.02M Tris HCl (pH 7.6).

Tris Buffered Saline – Tween (TBS-T)

0.8% (w/v) NaCl , 0.02M Tris HCl, 0.1% Tween (pH 7.6).

Tris EDTA (TE)

10mM Tris-HCl(pH8.0), 1mM EDTA. Autoclaved.

Triethylammonium acetate (TEAA) - (Transgenomics)

Tris-HCl

Stock solutions of 1M Tris base in dH_2O. Adjusted to desired pH with HCl. Autoclaved.

Trypsin Solution (Gibco-BRL)

(2.5%)

Vectashield® Mounting Medium with DAPI (Vector Laboratories)

Wave buffer A

0.1M Triethylammonium acetate (TEAA)

Wave buffer B

0.1M TEAA, 25% Acetonitrile

Wave buffer C

4.5% Acetonitrile

Wave buffer D

75% Acetonitrile

X-Gal

40mg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactosidase (X-Gal) made in dimethylformamide (DMF). -20°C. Kept in the dark.

2.2 Methods

2.2.1 Plasmid Preparation using the Promega Wizard Plus SV Miniprep

6 ml of LB supplemented with 100μg/ml ampicillin, was inoculated with the bacterial plasmid, and incubated overnight at 37°C. Glycerol stocks were then prepared for each culture by adding 250 μl of 80% Glycerol to 750 μl of culture, then mixing and storing the stock at -70°C. The remaining culture was then used for extraction of the plasmid as described by the manufacturer. DNA was stored at -20°C

2.2.2 Preparation of Mouse Genomic DNA

Method 1: For DNA extraction from tail tips and whole organs

The tail tip was cut 1cm from the end and added to 400 μl Tail buffer. 5-15μl of proteinase K solution (10mg/ml) was added to the sample. The amount of proteinase K added depended on the thickness of the tissue to be digested. Samples were shaken vigorously overnight at 55°C.

After the overnight digestion, 200μl saturated NaCl was added to each sample. The mixture was then gently inverted until it became homogeneously viscous and white. Samples were then centrifuged for 20 minutes at 10,000g. 400μl of supernatant was transferred into a fresh microfuge tube and added to 800μl 95% Ethanol. The tube was then inverted ten times before centrifuging at 10,000g for 1 minute. The supernatant was removed by pipette and 100μl 70% ethanol was added to the pellet.

The tube was re-centrifuged at 10,000g for 1 minute and the supernatant removed by pipette before allowing the pellets to air dry at 25°C for 30 minutes. DNA was resuspended in ddH_2O and quantified as described in 2.2.3. DNA samples were stored at 4°C until used, then archived at -20°C.

Method 2: For DNA extraction from ear-clips and toe-clips and embryonic tail tips

The tissue was added to 500 µl Tail buffer. 2.5µl of proteinase K solution (10mg/ml) was added and samples were shaken vigorously overnight at 55°C. Following digestion, the sample was inverted repeatedly until it became a homogenous mixture, then centrifuged at 10,000g for 2 minutes. The supernatant was then transferred to a clean microfuge tube, and 500 µl isopropanol was added and mixed with the sample by repeated inversion. The sample was then centrifuged for 10 minutes at 10,000g.

The supernatant was then removed by pipette, before allowing the pellets to air dry at 25°C for 30 minutes. DNA was resuspended in dH_2O and quantified as described in 2.2.3. DNA samples were stored at 4°C for use within 3 months, then archived at -20°C

2.2.3 Quantification of DNA Concentration

Spectrophotometry

DNA concentrations for low numbers of samples (fewer than 10), were determined by spectrophotometry using a Beckman DU640B spectrophotometer. DNA samples were added to water in UV cuvettes, mixed by pipetting, and placed in the spectrophotometer. Absorption was measured at 260nm and 280nm. Samples with a 260/280nm ratio below 1.8 potentially contained impurities in the solution so were repurified. The concentration of DNA was calculated using the absorption coefficient of 1 OD_{260}=50μg/ml of double stranded DNA.

For high numbers of samples (10 or more), a Spectramax 190 spectrophotometer (Molecular Devices) was used to measure DNA concentrations. DNA samples were diluted with dH_2O and mixed in 96-well, clear flat-bottomed plates (Corning). Analysis was carried out using SOFTmax® PRO 3.0 alias software (Molecular Devices), as described in the manufacturers instructions. Absorption was measured at 260nm and 280nm, impurity levels were calculated as with the Beckman DU640B spectrophotometer.

2.2.4 Oligonucleotide Primer Design and Synthesis

Oligonucleotide primers were designed using the Primer3 programme at the Whitehead institute website (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi), or through the local network. Primers for routine PCR were designed with a GC content of between 30 and 70%. Repeat sequences

and single base runs of more than 4 bases in length were avoided. Most primers were approximately 20bp in length, with annealing temperatures of at least 50°C.

Oligonucleotide primers were either ordered from external sources (MWG-Biotech, Genset, Sigma-Genosys), or were synthesized in house (by J.King) on a Beckman Oligo 1000M DNA synthesizer. Following Oligonucleotide synthesis, the base- and phosphorous protecting groups were removed to ensure the oligonucleotide was biologically active. The ultrafast Cleavage and Deprotection kit (Beckman) was used for the purpose, following the manufacturers instructions. Oligonucleotides were then passed through a NAP-5 column, and eluted in 1ml dH_2O. The concentration was determined by measuring the absorbency of a diluted aliquot at 260nm on a spectrophotometer, and assuming that for single stranded oligonucleotides, 1OD unit is equivalent to 20 μg/ml.

2.2.5 Polymerase Chain Reaction (PCR)

The polymerase chain reaction (PCR) was developed by Kary Mullis [206] and is used to exponentially replicate DNA at regions of interest. This process involves alternating temperatures to allow DNA melting, primer annealing, and DNA polymerase facilitated elongation of a novel complementary strand formed from dNTPs (Deoxynucleoside triphosphates).

PCR was performed on a MJ Research PTC-225 DNA engine tetrad thermo-cycler. They ranged from 10-100 μl in volume and were carried out in microtitre plates.

Each individual PCR contained 1 X PCR buffer, 0.2uM of each primer, 0.2-0.25mM dNTPs, 0.2-1 unit(s) of taq[5] and 5-50ng of template DNA. The yield of some reactions increased with the addition of either betaine (at a final concentration of 1M-1.5M), or Qiagen's Q solution (at a final percentage of 10% (v/v)). The plates were sealed with Easy Peel (ABgene), and placed on the thermocycler. One of the following programmes were then used, with the heated lid switched on:

Conditions for PCR were as follows: 200 uM dNTPs, 2.5mM MgCl2, 0.2uM forward and reverse primer, 0.45 U Applied Biosystems AmpliTaq Gold, and 0.025 U Pfu Turbo polymerase and 1X Ampli Taq Gold buffer. Reaction volumes were prepared in bulk and transferred into individual 96-well microtitre plate wells containing DNA pools (final volume 25-50 μl).

Programme 1 Standard PCR

(Genetic mapping, PCR screening, RT-PCR, multiplex PCR) with ABgene taq

95°C (15 minutes)[1]

35 cycles of 95°C (30 seconds), 55°C[2] (35 seconds), 72°C (45 seconds[3])

75°C (5 minutes)

Programme 2 Touchdown PCR

(PCR screening, Novel Marker Detection) with Hotstar taq, taq gold or Pfu Turbo polymerase.

95°C (15 minutes)[1]

9 cycles of 95°C (05 seconds), 55°C[2] (20 seconds) with a decrease of 1°C per cycle,72°C (45 seconds[3])

35 cycles of 95°C (05 seconds), 55°C[2] (20 seconds), 72°C (45 seconds[3])

Programme 3 DHPLC PCR

(Amplification of products for DHPLC analysis) with 0.45 U Taq Gold, 0.025 U Pfu Turbo polymerase and 1X Ampli Taq Gold buffer:

95°C (10 minutes) [1]

13 cycles of 94°C (20 seconds), 67°C[4] (1 minute) - with a decrease of 0.5°C per cycle, 72°C(1minute)

19 cycles of 94°C for 20 sec, 60°C[2] (1 minute), 72°C (1 minute)

72°C (5 minutes)

15°C hold.

Following amplification, heteroduplexes were formed with the following cycle conditions:

95°C for 4 min

45 cycles of 93.5°C for 1 min with a reduction of 1.5°C per cycle

25°C hold.

Programme 4 Site Directed Mutagenesis

(Mutagenesis of plasmids at specified loci) with Pfu Turbo polymerase.

95°C (3 minutes)[1]

12-16 cycles of 95°C (30 seconds),

58°C[2] (1 minute),

72°C (6 minutes and 30 seconds[3])

[1]HotStarTaq, AmpliTaq gold and Pfu Turbo DNA polymerases require an activation step. [2]The annealing temperature was changed depending on the T_m of the primers used in the reaction (ranging between 55 and 68°C). [3]The extension time was changed depending on the expected size of PCR product (ranging from 10 seconds to 2 minutes).[4]The initial annealing temperature for DHPLC product amplification should be 7°C higher than the annealing temperature for the second round of cycles.[5]DHPLC PCRs contained 0.45 U Applied Biosystems AmpliTaq Gold, and 0.025 U Pfu Turbo polymerase with 1X Ampli Taq Gold buffer

2.2.6 PCR Product purification

Primers, nucleotides, polymerases and salts were removed from PCR products, when necessary, using a Qiagen Qiaquick PCR purification kit. The DNA was eluted by adding 30 µl of dH_2O to the centre of the column and leaving to stand for 5 minutes before centrifuging the sample at 10,000g for 5 minutes. DNA concentration was determined as described in section 2.2.3.

2.2.7 Agarose Gel Electrophoresis of Nucleic Acids

Agarose gel electrophoresis was used to fractionate DNA using 1 X TBE buffer. PCR products were analysed on 0.8% - 2% agarose, depending on their size.

Agarose was added to 1X TAE buffer and heated in a microwave until boiling, this was then mixed and allowed to cool before adding 0.5 μg/ml ethidium bromide. The molten agarose was set in a gel plate with a 16 or 20-tooth comb on a flat surface. The gel was immersed in 1XTAE buffer; 0.5μg/ml ethidium bromide in an electrophoresis tank and the DNA samples, containing Orange G dye loaded into the wells. 1kb DNA ladder (Gibco-BRL) was used as a size standard.

Gels were electrophoresed at 50-150 V. The products were visualised under ultra-violet light (312nm) and photographed using AlphaImager2000 software for the visualisation of BAC DNA and novel marker detection, or AlphaImager2200v5.5 for all subsequent gels.

2.2.8 Extraction of DNA from Agarose Gels

A scalpel was used to excise DNA fragments from agarose gels. These samples were weighed, and the DNA extracted with a Qiaquick Gel Extraction kit (Qiagen). The DNA was eluted by adding 30 μl of dH_2O to the centre of the column and leaving to stand for 1 minutes before centrifuging the sample at 10,000g for 10 minutes. DNA was stored at -20°C.

2.2.9 Random Oligonucleotide Radiolabelling of DNA probes

A Megaprime labelling kit (Amersham) was used to label probes. Approximately 50ng DNA, in a total volume of 25 μl, was heated at 95°C with 5 μl labelling primers (random nonamer primers in dH_2O). These were then placed immediately on

ice for 5 minutes. The following constituents were then added: 10 µl labelling buffer (dATP, dGTP, dTTP in Tris/HCl pH 7.5; 2-mercaptoethanol and $MgCl_2$), and 2 µl (2 units) Klenow polymerase and 30-50 µCi [α-^{32}P]-dCTP (Amersham). The reaction mixture was incubated at 37°C for one hour.

Nick columns (Pharmacia) were used to separate unincorporated nucleotides from the radiolabelled probe. The volume of the probe was increased to 100 µl with dH_2O. After mixing, the probe was loaded onto a previously equilibrated Nick column (as described in the manufacturers instructions). Two volumes of 400µl dH_2O were added to the centre of the column and collected in two separate tubes. The second tube contained the radiolabelled probe.

Overgo End-Fill Method

All overgos used in this project were designed by Matthew Cadman. Overgos are overlapping oligonucleotides which can be used to amplify short sequence fragments. The overgos made to probe the RPCI-23 library were designed from the flanking sequence of known d13Mit STS markers.

A 10µl mixture of the forward and reverse overgos at 10µM were heated to 80°C for 5 minutes, then at 37°C for 10 minutes. 2µl of this mixture was then added to 1µl of klenow buffer, 3µl of a dA, dG and dT mixture (at 100µM), 1 µl klenow (2u/µl), 2µl H_2O and 1µl dCTP-^{32}P.

The reagents were mixed and incubated at room temperature for 1 hour before stopping the reaction with 1µl of 0.5M EDTA. The volume was made up to 100µl with ddH_2O then loaded onto previously equilibrated Nick columns (as described in the manufacturers instructions). Two volumes of 400µl dH_2O were added to the centre of the column and collected in two separate tubes. The second tube contained the radiolabelled probe. The probe was then denatured at 95°C for 5 minutes, placed on ice for 5 minutes, then added to the hybridisation buffer. Hybridisation was carried out as described under "Hybridisation of probes", but overgo probes were incubated overnight at 47°C and washed twice with 2XSSC/0.1%SDS for 15 minutes.

Blocking of Repeats in Probe with Mouse Cot-1

Mouse Cot-1 DNA (Invitrogen) was concentrated from 1mg/ml to 10mg/ml by ethanol precipitation. 25 µl 10mg/ml mouse Cot-1 DNA, 125 µl 20XSSC, and 5 µl 20%SDS were added to 400µl purified probe. This mixture was then heated to 95°C for 5 minutes in order to denature the DNA, then incubated at 65°C for 20 minutes before commencing with hybridisation

Hybridisation of Probes

BAC membranes were wetted with 0.2 X SSC before pre-hybridisation for 4-24 hours with Church Gilbert buffer at 65°C in rotating Hybridisation tubes. Each tube contained up to 4 BAC membranes, each membrane separated by a layer of gauze. The prehybridisation solution was then discarded. The radiolabelled probe was denatured at 95°C for 5 minutes before adding it to 25-40ml of prewarmed Church-Gilbert solution at 65°C. This mix was then added to the tube containing the membrane(s) and hybridised at 65°C for 14-20 hours. After hybridisation, the membranes were washed at 2 X SSC/0.1% SDS. Following this step, wash conditions were specific to each probe. 1XSSC/0.5% SDS and 0.5XSSC/0.5% SDS were used as necessary. Each wash was at 65°C for 30 minutes, the membranes were separated and incubated on shaking incubators when stringent washes were required. Membranes were wrapped in Saran wrap (Dow), and exposed to X-ray film in cassettes with intensifying screens at -70°C, from 5 hours to 1 week. The X-ray film was fixed and developed using the X-ograph Compact X4.

Genomic Library Screening by Hybridisation

2.2.10 RPCI23 Mouse BAC library.

The RPCI23 BAC library was constructed by K. Osoegawa, M. Tateno and P. de Jong (Roswell Park Cancer Institute, Buffalo), and was provided by the MRC HGMP-RC for members of the Mouse sequencing consortium. kidney and/or brain genomic DNA was extracted from Female C57BL/6J mice, then partially digested with EcoRI and EcoRI methylase. DNA was cloned into the pBACe3.6 vector at the EcoRI sites after size selection, and the ligation product transformed into DH10B cells. Clones averaged approximately 197kb in size. Clones were gridded onto 22.3cm^2 Hybond-N+ membranes at a coverage of 18,864 clones in duplicate, per membrane. 10 membranes were used for the complete library of approximately 170,000 clones, representing an 11.2-fold coverage of the mouse genome[207]

Primary Identification of Genomic Clones by Hybridisation.

STSs, cDNA probes and overgos and were used for genomic library screens. Pre-hybridisation, hybridisation and washing temperatures were 65°C with cDNA probes, and 47°C with overgo probes.

Either random priming or the overgo end-fill method (section 2.2.9) were used to label the probes. Any repetitive elements within the probes were blocked with mouse Cot-1 DNA. Overgo probes were pooled in batches of up to eight probes

together for genomic library screening. Hybridisation of the libraries was carried out as in section 2.2.9

Co-ordinates of positive clones were then determined according to the library manufacturers' instructions. Positive clones were ordered from the HGMP-RC for further characterisation (i.e.: secondary hybridisation and BAC-end sequencing).

2.2.11 BAC DNA Vector storage

BAC clones could be stored for up to one month in a host bacterial strain, at 4°C on LB agar plates containing the appropriate antibiotic. They could be stored indefinitely in 15% (v/v) glycerol at -70°C.

2.2.12 Restriction Enzyme Digestion of DNA

Restriction enzymes were used to digest DNA according to the conditions recommended by the manufacturer. DNA was typically digested in 1X restriction buffer and 1 enzyme unit/μg DNA (see table 2.1 for details of specific enzymes). Samples were incubated at the appropriate temperature for 1 to 16 hours before loading onto an agarose gel for electrophoresis as described in section 2.2.7.

List of Restriction Endonucleases

Enzyme	5'-3' Restriction site	Reaction temperature (°C)	BSA?	Buffer
Sfc I	5'-C^T Pu Py A G-3'	37	Y	NEB4
Bsr I	5'-A C T G G N^-3'	65	N	NEB3
Dde I	5'-C^T N A G-3'	37	N	NEB3
SnaB I	5'-T A C^G T A	37	Y	NEB4
Mlu I	5'-A^C G C G T	37	N	NEB3
Bgl II	5'-A^G A T C T	37	N	NEB3
Hind III	5'-A^A G C T T	37	N	NEB2
Kpn I	5'-G G T A C^C	37	Y	NEB2
Dpn I	5'-G A*^T C-3'	37	N	-
Sal I	5'-G^T C G A C-3'	37	Y	NEB3
Xba I	5'-T^C T A G A-3'	37	Y	NEB2
BamH I	5'-G^GATCC -3'	37	Y	NEB3

* = Methylated base only

^ = Position of cut on 5'-3' strand

Table 2.1: Restriction Enzymes used.

XbaI and BamHI were used for double digests as were BglII and MluI and the enzymes HindIII and KpnI. All enzymes in double digests were in buffers that allowed 100% efficiency, with the exception of KpnI which is 75% efficient in NEB2. DpnI was added directly to site directed mutagenesis PCR mixtures, following amplification of the respective target vectors.

2.2.13 Single Stranded Conformational Polymorphism (SSCP) Assays

Based on a method previously described [208]. SSCP gels were made using the following reagents: 1XTBE, 18% (v/v) acrylamide, 0.5% (v/v) bisacrylamide, 0.12% (w/v) ammonium persulfate (APS), 0.0004% TEMED (NNN'N' Tetramethylethyldiamine). Two glass gel plates, with spacers separating each pane, were aligned and gel mix was poured between them. A comb of 50 or 100 wells was inserted, and left for 1 hour at room temperature. Once set, the gel was clamped to a vertical electrophoresis tank containing 1XTBE buffer. DNA samples, containing 4 µl SSCP loading dye, were denatured at 95°C for 5 minutes, and placed immediately on ice. The samples were loaded into the wells. Electrophoresis was carried out at 36-38V for 8-16 hours at 4°C.

Following electrophoresis, the gel was removed from between the plates and placed in 10% ethanol, 0.5% acetic acid for 5 minutes, to fix the DNA. The gel was then stained with 0.1% (w/v) silver nitrate for 15 minutes. The gel was rinsed with dH_2O, and then developed in 1.5% (w/v) sodium hydroxide, 0.12% (v/v) formaldehyde until the bands were visible.

2.2.14 Novel Polymorphic marker Identification

DNA Preparation for Novel Marker Identification

Tail tips of 0.2-0.5mm were removed from mice of the following inbred strains: BALB/c, C3H/HeH, 101, C57BL/6, 3H1 (a C3H / 101 mosaic strain), Satin mice

(Foxq1 hair colour mutants on a 3H1 background). Tails were also taken from CAST/Ei and SPRET/Ei mice. DNA was extracted from this tissue as described in section 2.2.2.

Novel Microsatellite Identification

Sequence from individual clones were analysed by Joseph Weekes, Ann-Marie Mallon and Richard McKeone, using RepeatMasker (A.F.A. Smit and P. Green, unpubl.) (http://repeatmasker.genome.washington.edu/), to mask repeats from the sequence. The associated output file was then used to identify the positions of microsatellites (dinucleotides, trinucleotides, tetranucleotides and pentanucleotides). The microsatellite sequence and flanking regions were then screened with ePCR (http://www.ncbi.nlm.nih.gov/sutils/e-pcr/) to determine whether microsatellites at these loci had previously been identified and tested. If a sequence was not recognised by the database, primers were designed using sequence that flanked the microsatellite.

These primers were then used to amplify the microsatellite regions from the DNA of each of the mouse strains previously described, as described in section 2.2.5. Amplified DNA was then loaded onto agarose gels as described in section 2.2.7 and SSCP gels as described in section 2.2.13, and the band patterns were analysed for polymorhisms between strains.

Novel SNP Identification

Sequence from individual clones was analysed using PIPmaker (http://pipmaker.bx.psu.edu/pipmaker/) and also RepeatMasker. Regions within the deleted interval, with low sequence identity between mouse and rat and without repeats, were selected at random for SNP discovery. The selected sequences were used to design primers as described in section 2.2.4.

These primers were then used to amplify the selected regions from the DNA of each of the mouse strains previously described. Amplified DNA was loaded onto agarose gels (section 2.2.7) and SSCP gels. SSCP gels were prepared as described in section 2.2.13 with the following modifications to the protocol: Gels were prepared with and without 20mM Hepes incorporated into the gel mixture and buffer. Both types of gel were run at temperatures of 20°C and 4°C. This method has been previously described as improving the sensitivity of SSCP to conformational changes of DNA fragments [209].

2.2.15 Production of the Genotyping NED size standard

The NED size standard was produced in conjunction with Alison Hough, using the method of Lorraine Southam. Lambda DNA was added to a PCR that contained the forward primer and the appropriate reverse primer listed on table 2.3. Each PCR was carried out as described in section 2.2.5. Products above 100bp were gel purified as described in section 2.2.8. 30 µl of each purified product was then added to 60 µl of each unpurified product and 280 µl ddH_2O.

Primers Used In The Production Of The NED Labelled Genotyping Size Standard		
Primer Direction	Product Size	Primer Sequence
Unlabelled REVERSE primer (20pmol/µl)	70bp	5'-GTG AGA GAT AGA CTT TCT CCG TG
	80bp	5'-CCC GGA ATT TGT GAG AGA TA
	90bp	5'-GTT TAC CAG TCC CGG AAT TT
	100bp	5'-ACA GCG CCA TGT TTA CCA G-3'
	120bp	5'-GGA AAC AAT CGG CGA AAC GTA-3'
	140bp	5'-CGG GAA CGG ATA ACC TCA C-3'
	160bp	5'-ATT CAG AGG TGG AGC CGC-3'
	180bp	5'-CTT GAC ATA TCC CGG CGT A-3'
	190bp	5'-CAT GCT TCG GCT TGA CAT AT-3'
	200bp	5'-GGA TTC ACT TCA TGC TTC GG-3'
	220bp	5'-GGC GAC GCA GGG TCA TCT G-3'
	235bp	5'-GAT CTT CAT CCG GCA GGC-3'
	260bp	5'-TAA GCC GGG TCC GCC AGA-3'
	281bp	5'-ATG ATG ATG CGA CGG CGG-3'
	300bp	5'-TTC GTC ACG CAT GTT CTG CAT-3'
	320bp	5'-ACC TGA GCA ATG GCC AGC-3'
	340bp	5'-AAA CTG CCT GCA TCT CTT CG-3'
	360bp	5'-TTT GCC CTT AAG CAC GGC-3'
	380bp	5'-GCT TCA CCG GTC ATG GTG-3'
	400bp	5'-CCA CCT CAA CCG GAT CGA-3'
NED labelled FORWARD primer (17.7 pmol/µl)	-	5'-GAT CCG CTG TTT CTG CTG-3'

Table 2.2 Primers used for the NED size standard used for genotyping
Primers for the amplification of Lambda DNA that were used to produce a NED labelled size standard for genotyping

2.2.16 Genotyping

PCR products that demonstrated polymorphism between the test strains in section 2.2.14 were selected and sized by Ghazala Mirza of the Wellcome Trust Centre for Human Genetics. ROX, HEX and FAM Fluorescent dye labels were selected for these primer pairs by Ghazala Mirza, and products of a similar size (alleles with a size difference of less than 20bp) were labelled with different dyes in order to differentiate the products on analysis.

Tail tips of 0.2-0.5mm were removed from mice, by an animal technician in accordance with home office regulations. Tissue was stored at -20°C until processed. The DNA was extracted using a DNeasy® 96 Tissue Kit (Qiagen), using the protocol described by the manufacturer. Extracted DNA was then quantified using the Spectramax 190 spectrophotometer as described in section 2.2.3. DNA was diluted to give a working stock of 5ng/µl.

Amplification of pooled products for genotyping was carried out by PCR using Multiplex PCR kits (Qiagen) with an MJ Research PTC-225 DNA engine tetrad thermo-cycler. Reactions were carried out in microtiter plates in volumes of 12-15 µl. Each reaction contained 6µl of multiplex mastermix, 2-6 primer pairs (forward and reverse primers) at individual concentrations of 0.15 - 0.25 pmol/µl), and 10%(v/v) Q solution when amplifying primer sets which showed improved PCR efficiency in its presence.

Each plate of reactions was pooled in the following ratio: ROX products: 4 µl; HEX products: 3 µl; FAM products: 0.8 – 1.5 µl. 1 µl of this pooled mix was then added to 10 µl of formamide: NED size standard mix (50µl NED size standard - as shown in figure 2.2, in 1000µl formamide). The pooled samples were then loaded onto an ABI Prism 3130XL Genetic Analyzer (Applied Biosystems) according to the manufacturers instructions, using dye set D. Genescan® Analysis software version 3.1 (Applied Biosystems) was used for analysis of the raw data and calculation of the peak sizes for each dye used. This data was then downloaded into Genotyper (Applied Biosystems), which was set to identify specific alleles from peak size information.

Allelic size determination was repeated at Harwell and appropriate adjustments were made to the primer sequence in order to differentiate the products with the available equipment.

Dye colour	Marker multiplex pool	Number of products	Volume added to Final Pool (µl)
Green	HexA	4	3
	HexB	2	3
Blue	FAMA	4	1.5
	FAMB	4	0.8
Red	ROX	2	4

Table 2.3 Dye labels and volumes used

2.2.17 Ligation

When non-complementary sticky-ends underwent ligation, an initial end-filling stage was required before ligation could commence. DNA Polymerase I (Klenow) was added to the DNA, in the presence of 33 uM dNTPs and 1X NEB buffer2. Klenow was added at a concentration of 1u/μg of DNA. The reaction mixture was incubated at 25°C for 15 minutes, and stopped by the addition of EDTA to a final concentration of 10mM, and heating at 75°C for 20 minutes. End-filled products were then purified as described in section 2.2.6 or 2.2.7.

Ligations using T4 DNA ligase were performed in order to anneal blunt ends and sticky ends with their respective complementary sticky ends. In addition to the subject DNA, reactions typically contained 1X T4 DNA ligase buffer, and T4 DNA ligase in excess of 1 unit per 200 ng of template DNA. Reactions were incubated overnight at 16°C. Reactions were inactivated by heating the mix at 65°C for 10 minutes, then stored at -20°C until required

2.2.18 Transformation

Transformation using XL-1 blue supercompetent cells.

0.8 μl of β-mercaptoethanol was added to 50 μl of XL-1 blue supercompetent cells in an ice-cold sterile microfuge tube. The cells were then incubated on ice for 10 minutes, gently flicking the tube every 2 minutes to mix the cells. 1-5 μl of the plasmid DNA or ligation reaction was then added to the cells, depending on concentration of the plasmid. The tube was mixed by gentle flicking before placing

it on ice and incubating for **30 minutes,** then heat-shocked for **45 seconds** at 42°C. The sample was then immediately put on ice for 2 minutes, before adding **500µl NZY+ broth.** The sample was shaken at 220rpm for 1 hour at 37°C. The reaction mix was then spread onto 2 LB plates containing 100µg/ml ampicillin. These were incubated overnight at 37°C to grow visible colonies.

Confirmation of Transformants by Sequencing

Colonies were picked, using a sterile loop and prepared for wizard SV+ preparation as described in section 2.2.1. After completion of plasmid purification, plasmids were screened by restriction digestion (section 2.2.12) or sequencing (section 2.2.19).

2.2.19 Sequencing

All sequencing was performed with the dideoxynucleotide chain-termination method [210]. This technique utilizes labelled dideoxynucleotide triphospates (ddNTPs), unlabelled deoxyribonucleoside triphosphates (dNTPs), DNA template, an oligonucleotide primer and klenow, a DNA polymerase enzyme. The ddNTP molecules cannot form a phosphodiester bond with adjacent deoxynucleotides and they therefore terminate DNA chain elongation during polymerisation of the unlabelled dNTPs. Fragments are formed that begin at the primer sequence and terminate at any one of the potential nucleotides downstream of the primer. Each of the nucleotides downstream of the primer are represented as terminating ddNTPs of

differently sized fragments. Sequence is determined by analysis of the labelling signal produced by the terminating ddNTPs of size-ordered fragments.

The ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit was used to sequence all DNA templates using the appropriate primers.

Method 1

Template DNA was quantified by agarose gel electrophoresis (section 2.2.7) to determine the concentration and quality of the samples. Sequencing reactions were set up following the protocol of Dr. B. Roe, using one-twentieth the amount of ABI-recommended BigDye Terminator sequencing reaction mix (http://www.genome.ou.edu/big_dyes_plasmid.html). BigDye terminator sequencing reaction mix was diluted 1:2 with 5 X BigDye reaction buffer. A master mix was made for the total number of samples, containing the following components for each reaction: 2 µl BigDye diluted mix, 1 µl 10 mM primer, 100ng template DNA (up to 4 µl) in a 7 µl reaction volume, with the remaining volume made up with dH_2O. Cycle sequencing was performed on a MJ Research PTC-225 DNA engine tetrad, using the following conditions:

100 cycles of 95°C (10 seconds); 50°C (5 seconds); 60°C (4 minutes)

The sequence reactions were purified to remove unincorporated dye terminators from the samples, which can interfere with base calling. To each sequence reaction, 1 µl sodium acetate (pH4.6) and 18 µl 100% ethanol were added, transferred to a 0.2

ml tube and mixed by flicking the tube. The samples were placed on ice for 15 minutes before centrifugation at 10,000g for 30 minutes. The supernatant was discarded and the pellet washed with 100 µl 70% ethanol, and centrifuged for 5 minutes at 10,000g. The ethanol was removed and the pellet left to air dry for 5 minutes. The pellet was resuspended in 2 µl of loading buffer for ABI sequencing reactions, and the sample stored at -20°C prior to loading on the gel.

- Preparation of Denaturing Polyacrylamide Gel

The gel apparatus was set up, with two 36cm plates placed in a manifold, separated by 0.2mm spacers. The gel mix was made up fresh each time, with the following reagents: 18g urea, 25ml MilliQ water and 5.7ml 19:1 40% acrylamide:bis-acrylamide (BIORAD). The mix was stirred with Amberlite resin for 5 minutes to dissolve the urea. The solution was then degassed for 15 minutes. 27ml of gel mix were added to 3ml 10 X TBE (BDH), to which 150 µl 10% APS and 21 µl TEMED were added and mixed gently. The solution was poured between the glass plates using a 50ml syringe, tapping the plates continuously to avoid trapping air bubbles. The gel was left to polymerise at room temperature for a minimum of 2 hours.

- Electrophoresis on the ABI PRISM 377 Automated Sequencer

A Filter Set E matrix (instrument) file was made previously from the ABI PRISM dRhodamine matrix standards, according to the manufacturers instructions. The samples to be sequenced were denatured at 95°C for 5 minutes, and placed directly on ice prior to loading. Between 1.5 and 2.5 µl of sample was loaded into each well

of the polyacrylamide gel. Electrophoresis was carried out using run module "Seq Run 36E-1200", which was configured to read 36cm well-to-read gels, with 1200 scans per hour. Typically, electrophoresis was carried out for 7 hours. Data was collected and analysed by ABI PRISM DNA Sequencing Analysis Software, version 3.2

Method 2

2 µl BigDye v3.1 was used undiluted, and added to 2.5µl template DNA and 0.5µl primer at 20pmol/µl. Cycle sequencing was performed on a MJ Research PTC-225 DNA engine tetrad, using the following conditions:

100 cycles of 95°C (10 seconds); 50°C (5 seconds); 60°C (4 minutes)

In order to remove unincorporated dye terminators from the samples and to precipitate the purified product, the samples were treated with 15µl dH_2O, 2µl of 3M sodium acetate and 50µl of 95% ethanol. The reaction plate was sealed, gently shaken, and left in darkness at room temperature for 15 minutes. The samples were then centrifuged at 3000rpm for 45 minutes at 4°C. Following this, the plate was inverted onto a paper-towel and centrifuged at 1960rpm for 1 minute to remove the supernatant. The samples were stored at -20°C until sequencing.

Sequencing using the ABI Prism 3130XL Genetic Analyzer

No polyacrylamide gel was required for this machine. All sequencing that was carried out on the ABI Prism 3130XL was performed by Deb Brooker, Anne

Southwell and Zuzanna Lalanne, as described in the manufacturers instructions. The matrix used was Pop4, the run length (length of capillaries) was 36cm and the dye set, Z. Software used for the analysis of sequence data was Sequencing Analysis 3.7 (Applied Biosystems).

2.2.20 Total RNA Extraction from Mouse Eye and Lung Tissues

Mice were sacrificed and tissues were removed from the mouse. These tissues were immediately snap-frozen in liquid nitrogen and stored at -70°C. RNA was extracted from mouse tissue using a Qiagen RNeasy Mini Kit for mouse eyes and a Qiagen RNeasy Midi Kit for mouse lungs, according to the manufacturers instructions. A maximum of 250 mg of mouse tissue was processed at a time. Homogenisation was carried out using an IKA Labortechnik T25 basic homogeniser. The concentration of the total RNA was determined by spectrophotometry (section 2.2.3), given that an OD reading of 1 is equivalent to 40µg/ml RNA. The ratio of absorbance reading at 260nm and 280 nm (A_{260}/A_{280}) gave an indication of the purity of RNA. Pure RNA has an A_{260}/A_{280} ratio of 1.9-2.1 in 10mM Tris-HCl, pH7.5.

Additional RNA Quantification and quality control

An RNA 6000 Nano assay chip (Agilent technologies) was used in conjunction with an Agilent 2100 Bioanalyzer in order to determine concentration and quality of RNA samples, as described in the manufacturers instructions. Quality checks involved ensuring that peaks for 28S and 18S ribosomal bands were present, and that the 28S band was approximately twice the strength of the 18S band.

2.2.21 First Strand cDNA Preparation

First strand cDNA was prepared from 200ng-1μg of poly A+ mRNA with 1μl of oligonucleotide (either Oligo-dT at 50uM, gene specific primer at 2μM or random hexamers at 50ng/μl), 2.5μl of dNTPs at 10mM, 9.6μl Trehalose at 1.2M and DEPC H2O to make the reaction volume up to 25μl. The sample was incubated at 65°C for 5 minutes, and then placed on ice for 2 minutes. 5μl DDT at 0.1M, 5 μl of 10X RT buffer, 10μl MgCl2 at 25mM, 1μl RNase Out at 10u/μl and 200 units (1 μl) SUPERSCRIPT Reverse Transcriptase (GibcoBRL) were added and made up to 50μl with DEPC H2O before mixing and incubating at 55°C-65°C for 1 hour (preceded by a 5 minute incubation at 25°C when random hexamers were used). The reaction was inactivated by heating the mixture to 70°C for 15 minutes. cDNAs were stored at -20°C. Negative control reactions were also prepared as described above, but without the inclusion of reverse transcriptase.

2.2.22 Reverse Transcriptase PCR (RT-PCR)

5 μl of cDNA was used in RT-PCR assays. A 50 μl RT-PCR contained 1 X Pfu buffer (Stratagene), 0.3mM dNTPs, 0.4uM forward and reverse primers and 2.5 U Pfu Turbo (Stratagene). Genomic DNA and water controls were also included for each primer pair assayed. PCR conditions were as stated in programme 2 (section 2.2.5), with an annealing temperature of 55°C, and an extension time of 45 seconds. RT-PCR products were electrophoresed through an agarose matrix.

2.2.23 Bioinformatics

Manipulation of raw sequence was performed using ContigExpress, and DNA tools (http://biology.semo.edu/cgi-bin/dnatools.pl). Repetitive elements were identified through RepeatMasker, and were masked from the sequence for further analysis. Sequence similarity searches were performed against the various public databases using the BLAST suite of tools, through the National Center for Biotechnology Information (NCBI). ContigExpress was used for comparative analysis of sequence chromatograms. S.A.M was used by Mathew Cadman in the construction of the Del36H BAC contig.

A range of protein prediction programmes were utilized. PIX was hosted by the Human Genome Mapping Project Resource Centre (HGMP-RC) at www.hgmp.mrc.ac.uk/, which is no longer operational. Both nnpredict and NORSp were accessed via the PredictProtein–Meta portal (http://www.predictprotein.org/meta.php). Amino acid sequence comparisons of FOX genes were performed with the SIM alignment tool (http://expasy.org/tools/sim-prot.html). A database of protein structures was accessed to identify the structural characteristics of homologous regions in proteins related to FOXF2 (http://molvis.sdsc.edu/protexpl/frntdoor.htm).

Pairwise alignment of sequences was performed through the BLAST algorithm at NCBI. This site allows two sequences to be compared for the identification of regions of sequence similarity. Genomic sequences were also compared with

Pipmaker (http://pipmaker.bx.psu.edu/pipmaker/), which enabled the identification of non-conserved regions between two or more genomes.

2.2.24 F1 DNA and sperm archive

Two archives of DNA with parallel sperm archives were constructed at MRC Harwell by R. Cox, A. Hough (nee Hughill), D. Quwailid and co-workers [201, 202]. ENU mutagenised BALB/c males were crossed with C3H/HeH females and the DNA and sperm from male F1 offspring of this cross were archived.

The GSK-archive DNA consists of two sections of tail, whereas the new Harwell archive comprises whole tail, liver, kidney, spleen, and heart tissue. DNA for the GSK archive was extracted by A. Hough and DNA for the MRC archive was extracted by D Quwailid using Qiagen DNeasy 96 tissue kits. Test PCRs were performed on each DNA sample with primers from exon 2 of the Resistin gene (Retn), and these products visualised on an agarose gel as a further quality check. Sperm freezing was carried out as described by Nakagata and Takeshima [211] with modifications (Thornton et al. 1999) [212].

2.2.25 Denaturing High-Performance Liquid Chromatography (DHPLC).

Mutation detection was performed using a Transgenomic Wave machine that utilises DHPLC. The system was run according to the manufacturers instructions (Transgenomic). Primers for PCR amplification were designed in the introns at least

30bp from the splice sites of each exon (to allow for subsequent sequencing of the splice sites themselves) and with an optimal fragment length between 150 and 500bp.

DNA from four individuals were pooled using 12.5ng of each DNA giving a total of 50ng in the PCR. Sequence analysis was performed as described in section 2.2.19. Recovery by in vitro fertilisation was onto a C3H/HeH genetic background.

2.2.26 Pyrosequencer

SNP detection with the pyrosequencer required biotinylation of the forward primer. The reverse primer was left unlabelled. A sequencing primer was also designed by Zuzanna Lalanne, using software at the website. http://primerdesign.pyrosequencing.com/jsp/TemplateInput.jsp. Samples underwent PCR as described in 2.2.5 using the biotinylated forward and unlabelled reverse primers. Either Debra Brooker, Anne Southwell or Zuzanna Lalanne then loaded 20µl of the reaction mixture into a PSQ HS 96A instrument (Biotage) as described in the manufacturers instructions.

2.2.27 Mouse Necropsies

Eye Extractions for Histological Analysis

Eyes were enucleated from each mouse and fixed in 50% Karnovsky's medium for 45 minutes. This was followed by 3 X 30 minute washes in PBS before being stored

in PBS at 4°C until embedding and sectioning at Imperial College by the histology team there.

Lung Extractions for Histological Analysis

Once the lungs were removed, they were placed immediately in Neutral Buffered Formaldehyde and stored at 4°C until embedding.

Head Extractions For Palate Histological Analysis

After decapitation, heads were placed immediately in Neutral Buffered Formaldehyde and stored at 4°C until embedding.

Head Extractions For Eye Analysis

4% paraformaldehyde in PBS was heated to 65°C and mixed until the paraformaldehyde dissolved, then stored on ice. After decapitation, embryo heads were placed immediately into 10 ml of ice-cold fixative and left overnight at 4°C.

Following fixation, a series of 30 minute washes were performed in 10mls of solution, whilst gently shaking the sample. The wash steps were carried out as follows:

2 x 10 ml washes in PBS; 1 x 10ml 1:1 (v/v) PBS-ethanol mix; 2 x 10 ml 70% ethanol; 1 x 10ml 85% ethanol; 1 x 10ml 95% ethanol; 2 x 10 ml 100% ethanol.

The embryos were stored at 4°C in 100% ethanol until embedding at Imperial College.

Embedding and haematoxylin and eosin Staining

Embedding of mouse tissue in Paraffin wax and the subsequent haematoxylin and eosin staining was performed by the MRC MGU Histology unit for palate and lung analysis. Embedding was performed with a Shandom Embedding centre as recommended by the manufacturer. Haematoxylin and Eosin staining was carried out using a Shandon Varistain 24 device, according to the manufacturers instructions. Sections were taken using a Microm HM320.

2.2.28 Mouse Behavioural Studies

Mouse behavioural primary screening tests were carried out by the MRC Harwell Phenotyping core on a cohort of 32 littermate mice; 8 Foxf2 W-R mutant heterozygote females, 8 Foxf2 W-R mutant heterozygote males, 8 Foxf2 wildtype females, 8 Foxf2 wildtype males.

Protocols for the modified SHIRPA, Grip strength test, Y-maze test and swim test that were all performed on these individuals are all described in detail in the mouse phenotyping database, Empress (http://empress.har.mrc.ac.uk/browser/).

Light-Dark box: Mice were acclimatised for 5 minutes in individual cages attached to the Light-Dark box before they were allowed to explore the Truscan Light-Dark box arena (Coulbourn Instruments). The following parameters were automatically measured by the associated computer system: Latency, number of moves, time spent moving, velocity, distance covered, time spent in the light area, time spent in the dark area, number of times the light area was entered, vertical plane entries (rearing), central area entries in the light region, central area entries in the dark region.

Acoustic Startle & Pre-Pulse Inhibition

Startle box tests, which included "startle" and "prepulse inhibition" programmes were carried out as described in the Empress phenotyping document 10_005 (http://empress.har.mrc.ac.uk/browser/). Each test format incorporated pulses in a pseudorandom order, with the intervening period between each pulse not exceeding 25 seconds. PPI-1 and Startle-1 were used for primary screening, secondary screening was performed using all 4 of the programmes described below:

Startle 1

A 5 minute period of acclimatisation was followed by 10ms pulses at 110dB and at a frequency of either 12kHz, 20kHz or white noise. Mice were subjected to each pulse frequency a total of 5 times.

PPI 1

Mice were subjected to 10ms white noise pulses at 110dB which were preceded by 10ms pre-pulses of 90dB at either 12 kHz or 20 kHz. The pre-pulse was 90ms before the white noise pulse for each trial. There were 15 trials for each of the pre-pulse frequencies, and a further 18 trials that did not incorporate a pre-pulse

Startle 2

Mouse movement was measured 10 times with white noise 40ms pulses at 65dB to determine background motion levels. White noise 40ms pulses at 110dB were then used to test the mice 10 times with a delay of 20-30 seconds between trials.

PPI 2

Responses of mice to 10ms white noise at 70dB, 80dB, 85dB, 90dB and 110dB were tested 10 times for each volume. PPI trials incorporating further 10ms white noise pre-pulses at these volumes were performed 10 times at each volume, with a 40ms delay before 110dB white noise pulses lasting 40ms in each trial.

2.2.29 Mouse Metabolic Phenotype Studies

There was a reduction in numbers of the cohort for this aspect of the study: 4 Foxf2 W-R mutant heterozygote females, 5 Foxf2 W-R mutant heterozygote males, 7 wildtype females and 8 wildtype males were tested.

An Olympus AU400 Clinical Chemistry Analyser (Olympus Diagnostics, UK) and associated reagents were used by Tertius Hough for the Harwell Phenotyping Core. Mouse phlebotomy was carried out as described in Hough et al, 2002[213].

Metabolites tested with the Olympus AU400

- Sodium
- Potassium
- Chloride
- Urea
- Creatinine
- Total Calcium
- Phosphate
- Alkaline phosphatase (ALP)
- Alanine transaminase (ALT)
- Aspartate transaminase (AST)
- Total Protein
- Albumin
- Total Cholesterol
- High density lipoprotein cholesterol (HDL cholesterol)
- Triglycerides
- Glucose

The MRC Harwell Phenotyping Core also carried out Urinalysis and IPGTT (Intra-peritoneal Glucose Tolerance Test) as described at http://empress.har.mrc.ac.uk/EMPReSS/servlet/EMPReSS.Frameset.

2.2.30 Plasmid Preparation

Site Directed Mutagenesis

Primers were designed to a target region within a plasmid containing the gene of interest as described in the literature for the Quickchange® Site Directed Mutagenesis Kit. Forward and reverse primers included the required mutation in their approximate centre and were complementary to each other, spanning 25-40 bases and with a melting temperature (T_m) of greater than 78°C.

The supplied FOXF2 expression plasmid was used to transform Xl-1 Blue supercompetent cells (section 2.2.18) and subsequently harvested from these cells. PCR with the mutagenesis primers was performed to amplify the entire plasmid as described in section 2.2.5 programme 4, although the V410F mutagenesis primers required an annealing temperature of ~67°C. Each 50µl reaction mixture contained 5µl of plasmid at 20ng/µl, 125ng of each mutagenesis primer, 1µl Pfu Turbo polymerase, 4µl dNTPs at 1.25mM, 5µl pfu buffer (10X) and 67.75µmol betaine. Following thermal cycling, ten units of DpnI (0.5µl) was added to each reaction, mixed by pipette and incubated for 4 hours at 37°C.

Xl-1 Blue supercompetent cells were then transformed with this mixture, and the plasmid extracted and screened by sequencing as described in section 2.2.18.

Empty pEVRF0

A double digest of the FOXF2-pEVRF0 construct (500ng) was performed with 10 units of BamHI and 10 units of XbaI in order to remove the FOXF2 insert. After gel purification of the linearised plasmid, 1 unit of klenow was added to 200ng of the linearised pEVRF0 plasmid with dNTPs at 33uM in a 50µl reaction to allow blunt ended ligation as described in section 2.2.17. Following EDTA reaction termination, the product was purified with a Qiaquick PCR extraction kit.

4ng of this purified product was then added to a ligation reaction with 200 units of T4 ligase. Other conditions were as described in section 2.2.17.

These ligation products were used to transform XL1-Blue supercompetent cells. Single colonies were selected for harvesting, and were prepared using Wizard miniprep kits (as described in section 2.2.1), before screening the plasmids by direct sequencing with pEVRF0 primers that amplify the insertion site: CCCTGAAAACTTTGCCCCCTCC and GGGGGATCTTGGTGGCGTG.

Reporter Vector Construction

Oligonucleotide Multimerization

Two semi complementary oligonucleotides were designed that incorporated the human FOXF2 binding sequence that was determined by Pierrou and colleagues [205]. At the 5' end of each of these primers was a restriction sequence also present within the multiple cloning site of ptal-SEAP. The oligonucleotide containing the binding sequence as described by Pierrou carried an MluI restriction sequence at the 5' end and the oligonucleotide which contained the reverse complement of the previously described binding sequence carried a BglII site. Both oligonucleotides were synthesised with a phosphate group at their 5' ends.

1. **CGCGT**CCAAACGTAAACAATCCGA (**Bold** = MluI site. Underlined = FOXF2 binding site).
2. **GATCT**CGGATTGTTTACGTTTGGA (**Bold** = BglII site. Underlined = FOXF2 binding site complementary sequence)

A 20µl ligation reaction mixture was set up for multimerisation of the oligonucleotides. The reaction contained 2µl of the 10X ligation buffer, 8 units of T4 ligase, 150pmol of each multimerization oligonucleotide and was performed as described in section 2.2.17.

Reactions were incubated overnight at 16°C. They were inactivated by heating the mix at 65°C for 10 minutes and were then run on a 2% agarose gel (section 2.2.7)

from which DNA was extracted and purified (section 2.2.8) from band sizes consisting of multimers of X5, X7 and X9 double stranded oligonucleotide lengths (124, 172 and 220 bp respectively).

Meanwhile, the reporter vector pTAL-SEAP was linearised with a double digest of BglII and MluI before purification with a PCR purification kit.

The purified product was ligated to the oligonucleotide multimers with T4-ligase as described in section 2.2.17. The ligation product was used to transform XL1-Blue supercompetent cells as described in section 2.2.18. Colony PCR was performed on the resulting colonies with primers designed from either side of the pTAL-SEAP multiple cloning site (Forward: CAAAATAGGCTGTCCCCAGT; Reverse: CCTCCTCAACTGGGATGATG). Reaction conditions were as described in section 2.2.5, standard PCR but pipette tips were jabbed onto XL1-Blue colonies to remove a fraction of the colony for a DNA template. These pipette tips were twisted several times in order to mix the template into the reaction mixture before their disposal and the commencement of thermal cycling.

Long Oligonucleotides (~100bp) that contained multiple FOXF2 binding sites and therefore did not require multimerisation were also used for ligations of MluI/BglII linearised pTAL-SEAP vector. In common with the multimerisation primers, these oligonucleotides were synthesised with phosphate groups at their 5' ends. 0.064pmol of linearised pTAL-SEAP vector (20ng) was added to ligation reactions with 3X,

30X and 300X this molarity of the multiple binding site oligonucleotides, and 1.6u/µl of T4 ligase. Screening conditions were as described for the multimerisation oligonucleotides.

MluI Multiple binding site oligonucleotide:

CGCGTTTCCAAACGTAAACAATCCTTCCAAACGTAAACAATCCTTCCAAACGTAAACAATCCTTCCAAACGTAAACAATCCTTCCAAACGTAAACAATCCA

BglII Multiple binding site oligonucleotide:

GATCTGGATTGTTTACGTTTGGAAGGATTGTTTACGTTTGGAAGGATTGTTTACGTTTGGAAGGATTGTTTACGTTTGGAAGGATTGTTTACGTTTGGAAA

(**Bold** = Restriction sequence, underlined = FOXF2 binding sequence)

2.2.31 Probe Dig Labelling

The two multimerisation primers in section 2.2.30 which contained the human FOXF2 binding sequence were also utilised for the EMSA.

The oligonucleotide that included a BglII site and complementary sequence to the binding sequence described by Pierrou [205] (GATCTCGGATTGTTTACGTTTGGA) was labelled with digoxigenin-ddUTP

(Dig) in accordance with the protocol described in the Dig Oligonucleotide 3' – End Labelling Kit (Roche). Determination of labelling efficiency was performed as described in the manufacturer's instructions.

This oligonucleotide and its unlabelled semi-complementary counterpart (CGCGTCCAAACGTAAACAATCCGA) were then mixed at 50 fmol/µl, and incubated at 95°C for 10 minutes before cooling to 25°C to allow formation of the double stranded probe.

2.2.32 Cell Culture Techniques

Cos-7 cells were initially provided at passage number 7 (P7) by Richard Gale and Jane Baker.

Trypsinisation of Cells

Cos-7 cells were incubated at 37°C in T-75 flasks until a surface coverage of ~80% was obtained. Cell growth media was extracted before washing the culture twice with fresh DPBS. 1.5mls trypsin/EDTA was added to the cells (2.5mls for T-175 flasks) and the flask was rocked from side to side until all the cells had contact with the trypsin solution. The cells were then incubated at 37°C for 7 minutes before 10mls of fresh, pre-warmed media was added and mixed by pipette until the cell suspension was homogeneous.

Cos7 Cell Passage

Cell cultures were typically split upon reaching a surface coverage of ~80% to allow continued survival and for culture expansion when required. Following trypsinisation, a fraction of cell suspension was added to a new flask(s), depending on experimental requirements and based on the assumption that the cells approximately double their surface coverage during a 24 hour period.
Fresh pre-warmed growth medium was added to make the final volume in T-75 flasks up to 20mls and T-175 flasks up to 30mls, before incubating at 37°C.

Cryopreservation of Cells

Cells were trypsinised and 2 X 4.5mls of cell suspension was distributed into 2 X 15ml falcon tubes. These were centrifuged at 167 RCF (Relative centrifugal force) for 5 minutes before disposing of the supernatant

Both pellets were re-suspended and combined in 4mls of cryopreservation media (10%DMSO in Foetal Bovine Serum). This suspension was distributed into 4 X 1ml cryovial tubes. The cells were then placed in an isopropanol filled cryopreservation container, and stored at -70°C for 1 day before transferring the cryovials into liquid nitrogen for long-term storage.

Reviving Frozen Cos-7 Cell Stocks

Cryovials containing frozen Cos-7 cell stocks were removed from liquid nitrogen storage and thawed out at 25°C for 3 minutes, followed by 2 minutes at 37°C. The

thawed cell suspension was then added to 20mls of pre-warmed growth media. The media was then rocked from side to side in order to evenly distribute the cells before incubating at 37°C.

2.2.32a Optimisation of Transfection Conditions

Glass cover-slips were sterilised by immersion in 95% alcohol for 5 minutes. They were then allowed to air-dry within a Laminar air-flow hood for 10 minutes before placing them into 9.6cm^2 cell culture wells (Nunc).

Cos-7 Cells were trypsinised and diluted with fresh, pre-warmed (37°C) growth media to 1.75 X 10^5/ml before adding 1ml of this cell suspension to each well. The cells were mixed by rocking the wells from side to side, and incubated at 37°C for 24 hours.

Transfections of Cos-7 Cells were set up with the following reagent ratios, with the EGFP plasmid at 250ng/µl:

3µl genejuice, 1µg EGFP

1.5µl genejuice, 1µg EGFP

6µl genejuice, 1µg EGFP

3µl gnejuice, 0µg EGFP

3µl genejuice, 2µg EGFP

The genejuice was added to 97µl serum free DMEM media before vortexing then incubating the mixture at room temperature for 5 minutes. EGFP vector was then mixed into the media by gently pipetting up and down. Following a 15 minute incubation at room temperature, this transfection complex was dripped evenly across the media surface and the flask then rocked from side to side to allow an even distribution.

After 24 hours, the coverslips were washed with DPBS then the cells were fixed with 4% paraformaldehyde in FBS whilst shaking at 4°C for 20 minutes. The cells were washed twice with DPBS and placed on an orbital shaker at 4°C for 10 minutes. Coverslips were then placed on slides with 20µl vectashield & DAPI (Vector Laboratories). These were stored in the dark at 4°C until analysis of GFP transfection ratios with an Axioplan 2 microscope.

Blue filters were used to capture images of the DAPI stained cells (all cells) and green filters to visualise the GFP transfected cells. Photographs of 3 random fields of view at a magnification of X200 were taken for each coverslip with both the blue and green filters.

2.2.33 Transactivation Assay

Glass cover-slips were sterilised by immersion in 95% alcohol for 5 minutes. They were then allowed to air-dry within a Laminar air-flow hood for 10 minutes before placing them into 9.6cm^2 cell culture wells (Nunc).

Cos-7 Cells were trypsinised as described in section 2.2.32 and diluted with fresh, pre-warmed (37°C) growth media to 1.75 X 10^5/ml before adding 1ml of this cell suspension to each well. The cells were mixed by rocking the wells from side to side, and incubated at 37°C for 24 hours.

Transfections of Cos-7 Cells were set up using the genejuice manufacturer's instructions with the following reagent quantities:

97µl serum free media

4µl pooled vector DNA at a total concentration of 250ng/µl (450ng expression vector, 450ng pTAL SEAP (either with or without the FOXF2 binding site), 100ng pEGFP-C1 Transfection control.

3µl genejuice.

50µl growth media was extracted from just beneath the surface at 24, 48 and 72 hours after transfection of the cells. This was then centrifuged at 12,000 rcf for 10 seconds before aliquoting 2 X 18µl of the supernatant into separate tubes and storing at -20°C until analysis. A Great EscApe™ SEAP Detection Kit was used for the determination of secreted alkaline phosphatase quantities in these media samples as described in section 2.2.34

The coverslips were then treated and analysed as described in section 2.2.32a.

2.2.34 Reporter Activity Quantification

50µl growth media was extracted from just beneath the surface. This media was centrifuged at 12,000 rcf for 10 seconds before aliquoting 2 X 18µl of the supernatant into separate tubes and stored at -20°C until analysis by either the Great EscAPe SEAP Chemiluminescence Detection Kit or the MightyLight™ Rluc Assay kit, depending on the transfection control used.

Secreted Alkaline Phosphotase (SEAP) Quantification

Levels of SEAP were quantified using a Great EscAPe SEAP Chemiluminescence Detection Kit. Samples were subjected to the procedure described in the manufacturers instructions, alongside a positive control placental alkaline phosphatase (BD Biosciences) at concentrations of 1000,000 pg/ml; 200,000 pg/ml; 40,000 pg/ml; 8,000 pg/ml; 1,600 pg/ml; 320 pg/ml; 64 pg/ml; and 12 pg/ml. The luminescence of the standards and samples was detected at 20 and 40 minutes after the addition of CSPD with a Luminoskan Ascent luminometer (Thermo Scientific).

Luciferase Quantification

Samples which had been transfected with the pMluc-3 CMV vector as a transfection control required the detection of luciferase levels within the media. The MightyLight™ Rluc Assay kit was used as described in the manufacturers instructions in the Sample preparation – Culture Medium section and Assay procedure section.

2.2.35 EMSA

Protein Extraction

Cell cultures were trypsinised and resuspended as described in section 2.2.32. A haemocytometer was used to estimate cell density

The volume of media equivalent to 5.25 $X10^6$ cells was added to each T175 flask, for an estimated cell count of 1.5 $X10^7$ at vector transfection, 24 hours after seeding.

Transfections were performed as described in section 2.2.32, with the following reagent volumes:

Genejuice: 56µl

Total Vector: 70µl at 250ng/µl (17.5µg)

Expression Vector: 63µl at 250ng/µl

pTAL-SEAP or pMLuc-3 CMV Transfection control vector: 7µl at 250ng/µl (1750ng)

Serum free media: 2200µl

48 hours after these transfections, extraction of media for the analysis of reporter activity was performed with either the Great EscAPe SEAP Chemiluminescence Detection Kit or the MightyLight™ Rluc Assay kit (section 2.2.34).

Immediately following the removal of the media, the transfected cells were trypsinised (section 2.2.32) before the nuclear and cytoplasmic fractions were

extracted as described in the Nucbuster protocol. These fractions were each then subdivided into 10 separate aliquots before storage at -70°C.

Protein Quantification

Protein quantification was carried out with a BioRad Dc Protein Assay kit as described in the manufacturers instructions, with the following reagent quantities:

Sample or standard: 25µl

Reagent A': 125µl

Reagent B: 1000µl

BSA was used for the protein standard at 20 µg/µl, 10 µg/µl, 2 µg/µl, 1 µg/µl, 0.5 µg/µl, 0.25µg/µl and 0.125 µg/µl.

EMSA Reaction and Analysis

The DNA binding reaction described in the Nucbuster protocol was performed with the following reagent quantities:

4 X Binding buffer: 5µl

CotI DNA or Salmon sperm DNA: 1µl at 50ng/µl

Poly (dI-dC): 0.01 - 0.03u (1-3µl at 0.01u/µl)

Nuclear extract: 15, 30 or 40µg

ddH2O to 17µl

These reagents were mixed by pipette before adding 1µl of the Dig labelled probe at 50fmol, mixing, and incubating on ice for 30 minutes. The binding mixture was then either loaded onto a protein gel immediately or 1µl FOXF2 antibody (either FOXF2 antibody 1 (Abcam) or FOXF2 antibody 2 (Abnova)) was added and the mixture was incubated on ice for a further 20 minutes before loading onto the gel for a supershift assay. Proteinase inhibitor was added to antibody 2 for some of the supershift runs. Unlabelled probe was also added as a specific inhibitor at 1X-100X the concentration of the labelled probe. Following this reaction, the binding mixture was loaded onto a native polyacrylamide gel.

Two gel systems were used in order to optimise visualisation conditions.

1. Bio-Rad Mini-Gel System. Recipes of 5% and 6% were used:

5% Polyacrylamide Gel

40 % polyacrylamide (29:1): 1250µl

10 X TBE: 500µl

ddH2O: 8135µl

Temed: 15µl

APS: 100µl

6% Polyacrylamide Gel

40 % polyacrylamide (29:1): 1500µl

10 X TBE: 500µl

ddH2O: 7885µl

Temed: 15µl

APS: 100µl

2. CriterionTM 5%TBE pre-cast gels (Bio-Rad)

Gels were pre-run at 4°C for 30 minutes at 100v in 0.5 X TBE. Samples were then loaded and the gel was run for an additional 1 hour. Gels were blotted onto Hybond N+ membrane overnight at 4°C in 0.5 X TBE at 30v. Membranes were then cross-linked with a UV Stratalinker (Stratagene).

A DIG Wash and Block Buffer Set (Roche Diagnostics GmbH) was used to develop EMSA membranes as described in the manufacturer's instructions. CSPD was added to the membrane which was then wrapped in Saran wrap (Dow) before incubating at 37°C for 10 minutes. The wrapped membrane was then exposed to chemiluminescence film in cassettes with intensifying screens at 25°C, from 20 minutes to 1 hour. The film was fixed and developed using an X-ograph Compact X4. Band strengths were quantified using a TD2 Transmission Densitometer.

2.2.36 Antibody Detection for FOXF2 transfection confirmation

Protein from each FOXF2 transfection and a HepG2 control were added to 2.5µl NuPage LDS sample buffer, 1µl Nupage reducing agent and H_2O to make the sample up to 6.5µl before heating to 70°C for 10 minutes. These samples were then loaded onto a Nupage® Novex Bis-Tris gel and run with XCell Surelock™ Mini-Cell apparatus at 200v in 1XSDS Running buffer with 0.25% NuPage® antioxidant. Blotting was carried out at 30v for 1 hour onto 0.45 µm nitrocellulose membrane (Invitrogen) as described in the gel manufacturers instructions.

After blocking the non-specific binding sites with either 5% powdered milk or 5% BSA for 1 hour on an orbital shaker, the membranes were washed for 5 minutes in PBS-T or TBS-T containing
0%-1.5% blocking reagent (wash solution).

FOXF2 antibody 1 (Abcam) was added to PBS-T or TBS-T with 1.2%-1.5% blocking reagent, to a concentration of 0.5-5µg/ml. This antibody solution was added to the membrane and incubated at room temperature on an orbital shaker for 1 hour. Following a 15 minute wash and two 5 minute washes in wash solution, the membrane was incubated in a 1:75,000 dilution of the secondary antibody, Anti-Rabbit IgG peroxidise, in PBS-T or TBS-T with 1.2%-1.5% blocking reagent.

The membrane was washed on an orbital shaker once for 15 minutes and four times for 5 minutes in wash solution before developing the membrane. 500µl of ECL

solution 1 and 500µl ECL solution 2 which were supplied with the ECL™ Western Blotting Analysis System were mixed after allowing both reagents to equilibrate to room temperature. This mixture was added to the surface of the membrane and incubated at room temperature for 1 minute before draining the excess fluid and wrapping the membrane in Saran wrap. The wrapped membrane was then exposed to chemiluminescence film in cassettes with intensifying screens at 25°C, from 2-10 minutes. The film was fixed and developed using an X-ograph Compact X4.

Chapter 3: Mapping the Region of Mouse Chromosome 13 Deleted in the Del(13)Svea36H Mutation

3.1 Introduction

The Del(13)Svea36H deleted region (henceforth known as the Del36H region) was detected at Harwell during an X-ray mutagenesis screen[82] . The deletion causes early embryonic lethality in homozygote carriers, indicating the presence of developmentally significant genes within the region. A recessive coat-colour allele (the *Satin* allele of *Foxq1*) located within the deleted interval has allowed the use of this deletion as a tool for the detection of lethal recessive mutations. The mutagenised genes that cause the lethal phenotype can then be identified and their function investigated [26].

This region of the mouse genome shows conserved synteny with human chromosome 6p22.1– 6p22.3 and 6p25. Large sections of these two genomic regions have been deleted in 6p deletion syndromes, which consist of distinct but overlapping deletions. Symptoms of these syndromes include ocular dysgenesis, orofacial clefting, deafness, and brain, heart and kidney defects; although the range of symptoms varies widely depending on the extent of the respective deletion. There are also several specific diseases which map to the region such as Fanconi's anaemia [214], haemochromatosis [215], schizophrenia [216], Behcet's disease [217] and cervical carcinoma [218, 219]. A characterisation of mutations within the mouse Del36H deletion may help to shed light on the genes involved in both the 6p deletion syndromes and other diseases caused by mutations within the chromosome 6p region.

The construction of a sequence ready, annotated map was initiated in order to facilitate the mapping of mutations arising from the lethal recessive phenotype screen. The Del36H region spans 12.7Mb of proximal mouse chromosome 13 and is relatively gene-rich (201 genes) in comparison with the genome average [84] Original estimates put the size of the region at approximately ~25Mb based on cytogenetic data and genetic maps derived from the EUCIB resource. This is significantly larger than the 12.7Mb observed in the physical map, but because the physical map is now fully contiguous, this 12.7Mb calculation is the more reliable figure. Potential reasons for the apparent discrepancy between the size predicted by the sequence and genetic maps may involve recombination hotspots within the interval.

Markers within the region had been placed on radiation hybrid maps [21, 36], allowing their incorporation into genetic and YAC maps [37, 220] as shown in figure 3.1. The STS data accrued during the construction of these maps provided a framework for subsequent bacterial clone-based physical maps. This region was selected by the MRC UK mouse sequencing consortium as one of the sites to be targeted for a hierarchical, clone based sequencing approach, in addition to the whole genome sequencing approach that was used for sequencing the majority of the mouse genome. PACs from the RPCI-21 library were used in preliminary stages of the physical map construction, but the larger size of BAC inserts (in addition to their increased stability compared with YAC clone inserts) meant that BACs were

selected as the best clone system available at the time in terms of efficiency and data reliability.

The clone based approach has the advantage over the whole genome shotgun approach in that it is less likely to result in the loss of sequence data due to the compression of regions with a high level of sequence identity [43] [221].

A panel of informative markers has been identified for mapping of mutant lines arising from the lethal recessive screen. This has been used to reduce the critical regions of 11 mutants

3.2 Physical Mapping of Del36H

An array of markers were selected from the estimated Del36H region [184, 220, 222] including 24 MIT/Whitehead markers [35], 10 genic probes, 50 ESTs (Expressed Sequence Tags) and 8 human homologous STS (Sequence Tagged Sites). These markers were used to probe RPCI-23 Mouse genomic BAC libraries (provided by MRC HGMP-RC for members of the Mouse sequencing consortium). The majority (155) of these 176 probes were pooled in batches of 5-10 when performing the hybridisations although 57 hybridisations were performed with single probes (32 probes were hybridised both singly and in pools). Pool sets are listed in table 3.1.

The MIT markers were not used to probe the BAC libraries directly because the microsatellite regions within these sequences would have resulted in hybridisations

to multiple genomic loci. Instead, the known flanking sequence was used by M. Cadman to design overlapping oligonucleotides (overgos). These overgos were incubated with a Klenow polymerase and dNTPs including ^{32}P labelled cytosine. This formed a double stranded probe which was labelled on one strand with ^{32}P which was then denatured before exposure to the BAC library for hybridisation.

3.3 Source of Probes used for Hybridisation of the RPCI-23 BAC Library

The probes which were used to screen the RPCI-23 BAC library were from a wide range of sources, a majority of them had previously been used for radiation hybrid and genetic mapping of the deleted interval [21]. Many of the probe sequences originated from the MIT/Whitehead institute, including the D13Mit markers, random STS markers and the M-prefixed ESTs. Many of these had been localised to proximal chromosome 13 by radiation hybrid mapping at the MIT/Whitehead institute, whilst others were previously mapped by YAC contigs anchored to chromosome 13.

The Cwr markers are STSs which were identified by Hong and Chakravarti et al [184] during positional cloning of the ch (congenital hydrocephalus) phenotype, which they localised to *Foxc1* (then known as mf1).

ESTs which had previously been mapped to the region by the MRC UK Mouse Genome Centre and the EEC Mouse Transcript Mapping Consortium were also used.

Pools of Probes Used to Screen the RPCI-23 BAC Library						
FS1	**FS2**	**FS3**	**FS4**	**FS5**	**FS6**	**FS7**
D13Mit294	M-08497	Hfe3	M-01319	Btn	M151-C02	M176-C02
D13Mit265	M-09785	M141-B07	M-07528	366h7t7	M161-G05	AA407021
D13Mit307	M-131-E02	M-11510	M135-A07	525F18t7	M199-H07	Ai115341
cwr11	M-11527	AI225904	M122-d09	584b16t7	M203-D06	Sox4_1
cwr6	M-142-H01	M153-A03	M-09543	Cwr1	M211-E01	Sox4_4
cwr10		K01661		Cwr3	M212-E04	Olfr11
Mf1				Cwr31	M220-E01	D13Mit79
M-156-d11				Cwr32	AA516957	K00020
M-09825					AA571384	K00153
M-04234					R 74621	K03922
M-08353						
M-01316						
FS8	**FS9**	**FS10**	**FS11**	**FS12**	**FS13**	**FS14**
K00879	K00181	M257-E09	M-00890	HB001	39M12sp6	72c13t7
K01931	K00829	M275-F01	M-08279	HB002	290i8t7	51a11sp6
K02035	K01193	M280-d05	M-10786	HB003	468M24sp6	94b10t7
16F12t7	80c18sp6	M301-H08	M-08298	HB004	Y12294	335h9t7
16F12sp6	87c10sp6	M239-C02	M-08493	HB005		241e24sp6
18F15 t7	92g13t7	M265-G02		HB007		K03211
18F15sp6	92g13sp6	M253-H12		27E3sp6		K03265
34M15sp6	278e16t7	M237-F06		11F17sp6		M-08270
34M15t7	278e16sp6	M284-F11		414k13sp6		
80c18t7	298j12sp6	M285-C08		404N10t7		
				467F12t7		
				71F18t7		
				439F14sp6		
FS15	**FS16**	**FS17a**	**FS17b**	**FS18**	**FS19**	**FS20**
M203-B05	Cwr26	12F15t7	271B10sp6	D13Mit197	D13Mit306	K01931
M289-B08	Cwr27	312d21sp6	266F10t7	D13Mit273	D13Mit15	K02035
K03576	343e12sp6	310c7t7	326c24sp6	D13Mit133		
	423H18sp6	310c7sp6	408j22t7	D13Mit220		
	490K5sp6	112o16t7	254H12sp6	D13Mit15		
	509H17t7	254H12sp6	HB008	D13Mit84		
	Cwr12	349k3sp6				
	Cwr13	148F21t7				
	392L15sp6	122c3sp6				
	Cwr30					

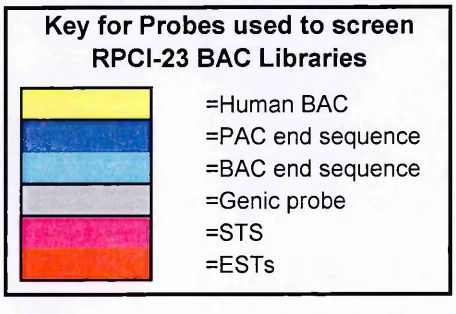


Table 3.1: Pooled hybridisation probes
The pooled probes which were hybridised to the RPCI-23 BAC library.

Single Probes Used to Screen the RPCI-23 BAC Library							
D13Mit58	D13Mit133	D13Mit136	D13Mit241	D13Mit265	D13Mit294	D13Mit307	D13Mit206
D13Mit81	D13Mit154	D13Mit82	D13Mit86	D13Mit83	D13Mit273	D13Mit84	D13Mit116
D13Mit175	D13Mit220	D13Mit135	D13Mit85	D13Mit16	D13Mit14	D13Mit15	D13Mit17
D13Mit197	D13Mit163	D13Mit176	M-09785	M-11691	M-09543	M-11510	M-07528
M-05084	M-01319	M-11527	M-08497	M-00442	M-09825	M-05614	M-01316
M-08353	M-04245	M-04234	Cwr6	Cwr10	Cwr11		
M-122-DO9	M-135-A07	M-153-A03	M-142-H01	M-131-E02	M-156-D11	A1225904	K01661
Hfe3	Gpx5 Int1	Spi3					

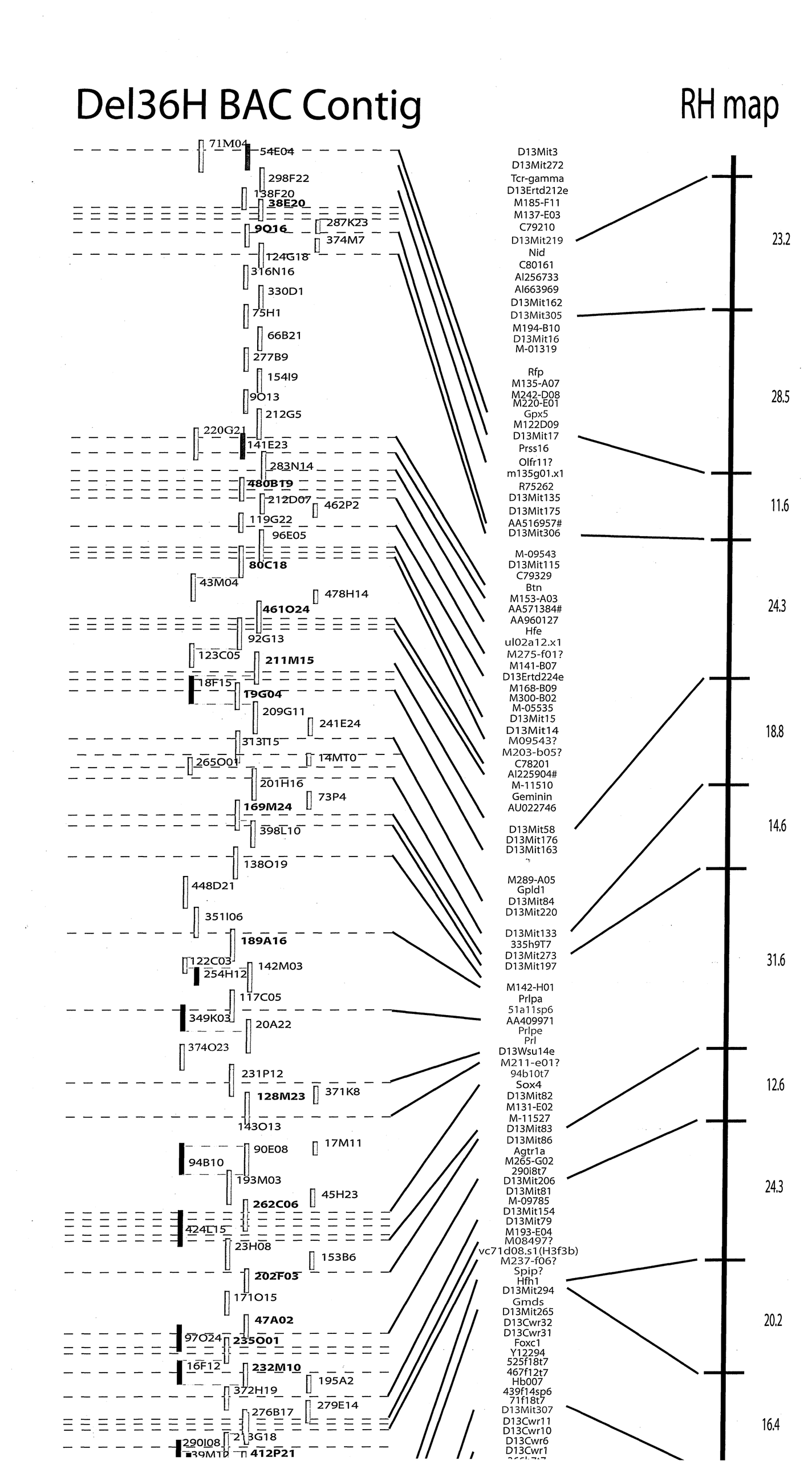
Table 3.2: Single hybridisation probes
List of the single probe hybridisations performed against the RPCI-23 BAC Library. The colour code is as shown in table 3.1 (multiple probe hybridisations).

3.4 Construction of the Physical Map

The single probes outlined in table 3.2 were hybridised to the RPCI-23 library membranes as outlined in section 2.2.9. Clones to which these probes hybridised were sequenced at their T7 and Sp6 sites by C. Sellick, in order to obtain end sequence data which could be used for a chromosome walking approach in contig construction. Probes designed from this sequence were pooled into batches of 3-13 probes for an increased throughput rate (table 3.1). Probes which had been mapped to similar regions were pooled together where possible, but probes were usually not fine-mapped so were pooled at random.

Singleton data from this screening is shown in appendix 3.4. These were clones that hybridised to only a single probe or probe pool. They couldn't be used in the construction of the physical map as an approximation of the orientation, position and size of a clone within a contig is impossible without multiple hits. A second hit on a clone is also further evidence that the clone is not a false positive.

The hybridisation data were combined with fingerprint data from the RPCI-23 and RPCI-24 BAC libraries[223] by Matthew Cadman, who used the output to construct a physical map of the region [26] – (figure 3.1). A BAC coverage of at least 9-12 fold was required before selecting clones for sequencing. This was in order to minimise the probability of the selection of chimaeras or unstable clones and to reduce the potential for sequence redundancy. The physical map of Del36H now consists of a fully contiguous minimal tiling path (MTP) that was submitted to the Sanger centre for shotgun sequencing [224]. This was to contribute towards the eventual construction of the mouse genome sequence.


Del36H BAC Contig
RH map
71M04
54E04
298F22
138F20
38E20
9Q16
287K23
374M7
124G18
316N16
330D1
75H1
66B21
277B9
154I9
9O13
212G5
220G21
141E23
283N14
480B19
212D07
462P2
119G22
96E05
80C18
43M04
478H14
461O24
92G13
123C05
211M15
18F15
19G04
209G11
241E24
313I15
265O01
14MT0
201H16
73P4
169M24
398L10
138O19
448D21
351I06
189A16
122C03
254H12
142M03
117C05
349K03
20A22
374O23
231P12
128M23
371K8
143O13
90E08
17M11
94B10
193M03
45H23
262C06
424L15
23H08
153B6
202F03
171O15
47A02
97O24
235O01
16F12
232M10
195A2
372H19
279E14
276B17
213G18
290I08
412P21
D13Mit3
D13Mit272
Tcr-gamma
D13Ertd212e
M185-F11
M137-E03
C79210
D13Mit219
Nid
C80161
AI256733
AI663969
D13Mit162
D13Mit305
M194-B10
D13Mit16
M-01319
Rfp
M135-A07
M242-D08
M220-E01
Gpx5
M122D09
D13Mit17
Prss16
Olfr11?
m135g01.x1
R75262
D13Mit135
D13Mit175
AA516957#
D13Mit306
M-09543
D13Mit115
C79329
Btn
M153-A03
AA571384#
AA960127
Hfe
ul02a12.x1
M275-f01?
M141-B07
D13Ertd224e
M168-B09
M300-B02
M-05535
D13Mit15
D13Mit14
M09543?
M203-b05?
C78201
AI225904#
M-11510
Geminin
AU022746
D13Mit58
D13Mit176
D13Mit163
M289-A05
Gpld1
D13Mit84
D13Mit220
D13Mit133
335h9T7
D13Mit273
D13Mit197
M142-H01
Prlpa
51a11sp6
AA409971
Prlpe
Prl
D13Wsu14e
M211-e01?
94b10t7
Sox4
D13Mit82
M131-E02
M-11527
D13Mit83
D13Mit86
Agtr1a
M265-G02
290i8t7
D13Mit206
D13Mit81
M-09785
D13Mit154
D13Mit79
M193-E04
M08497?
vc71d08.s1(H3f3b)
M237-f06?
Spip?
Hfh1
D13Mit294
Gmds
D13Mit265
D13Cwr32
D13Cwr31
Foxc1
Y12294
525f18t7
467f12t7
Hb007
439f14sp6
71f18t7
D13Mit307
D13Cwr11
D13Cwr10
D13Cwr6
D13Cwr1
23.2
28.5
11.6
24.3
18.8
14.6
31.6
12.6
24.3
20.2
16.4

3.5 Lethal Recessive Screen

The Early Development group at Harwell utilised the Del36H deletion for a regional lethal recessive screen. *Satin*, a recessive coat colour allele of *Foxq1* was used to highlight potential lethal mutations. Homozygous *Satin* males were mutagenised with ENU and their male offspring were mated to del36 deletion carriers. Progeny resulting from this cross that exhibited a *Satin* coat would be hemizygous deletion carriers which carried the *Satin* allele (see figure 3.2). A reduction in the numbers of *Satin* coated individuals resulting from such a cross would suggest linkage between Foxq1 and a hemizygously lethal mutation.

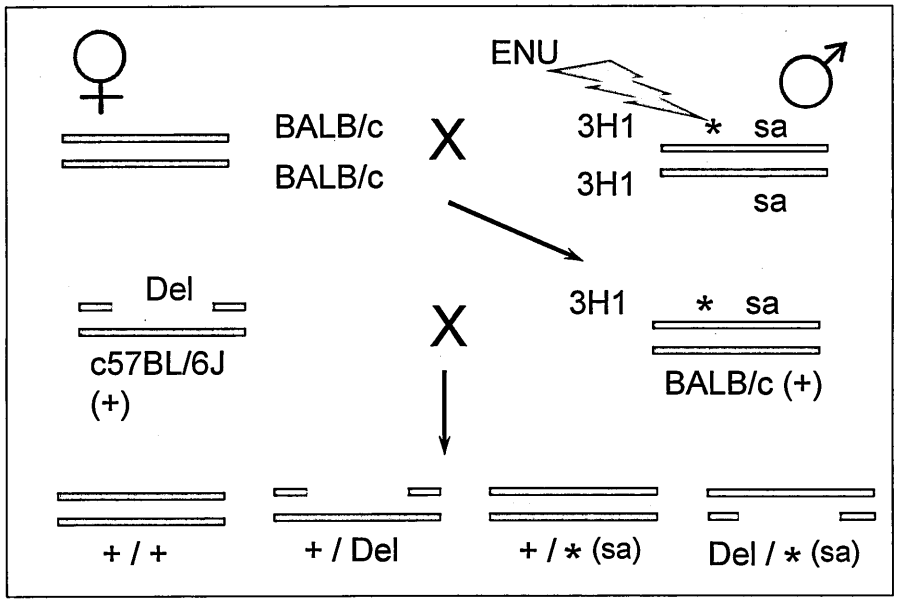


Figure 3.2: Breeding Scheme for the Lethal Recessive Screen
Following ENU injection, *Satin* coated C3101F1 (3H1) males were bred with BALB/c females until 30 male progeny were available for further breeding. Male offspring were then mated with female Del36H deletion carriers until at least 25 progeny were obtained from each G1 cross. Those lines which produced a significantly lower than expected ratio of *Satin* coated individuals (Del/*(*sa*)) were considered as potential lethal mutations and underwent inheritance testing for evidence of the heritability of the mutation.

3.6 Mouse Strains Used for the Lethal Recessive Screen

Four mouse strains were involved in the lethal recessive screen as shown in figure 3.2. The F1 hybrid progeny of C3H/HeH (henceforth known as C3H) and 101/H have been referred to as 3H1 throughout this work.

ENU mutagenesis was performed on males that were homozygous for the *Satin* allele. The satin colony was maintained as a closed colony segregating C3H and 101 alleles

The two mouse strains BALB/c and C57BL/6J (henceforth known as B6) were used at various stages of outcrossing during mutation testing and recombinant mapping of mutant lines (see sections 3.5 and 3.12 for further details).

3.7 Microsatellite Marker Design and Polymorphism Testing

A dense STS map was built up during and prior to the construction of the minimal tiling path contig as described in sections 3.1-3.4. Despite this, however, there was a dearth of available polymorphic markers for fine mapping of mutant loci. Markers were required within the Del36H region that were polymorphic between the strains involved in the lethal recessive screen (figure 3.2). These were intended for recombinant mapping for the localisation of mutations arising from the screen. Microsatellites have a tendency to undergo 'slippage' on replication [225] causing variability in the size of their alleles, making them ideal loci to screen for polymorphisms between inbred mouse strains.

Novel microsatellite markers were generated within the Del36H region using draft sequencing data from the BACs selected for the sequencing of the Del36H interval (section 3.4).

BAC sequence was screened for microsatellite repeats using web-based repeat masker software [220]. Primers were designed flanking these repeats, which were then used to amplify loci in each of the strains involved in the recessive screen in addition to mice with a 3H1 background and homozygous carriers of the *satin* allele (B6; BALB/c; C3H; 101/H; 3H1; *satin*). SSCP and agarose gel analysis was then performed on each sample to identify polymorphisms between the strains as described in sections 2.2.13 & 2.2.14.

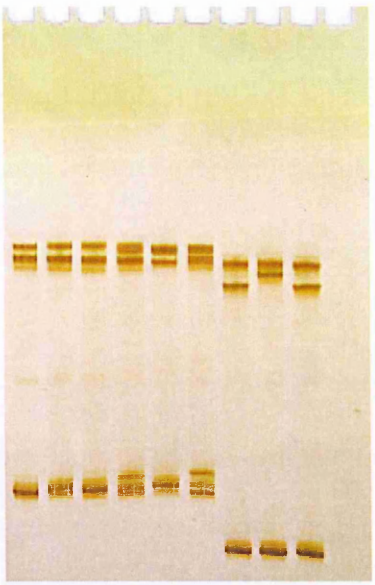


Figure 3.3: An SSCP gel
This technique involved adding denatured PCR products to the wells of a 0.5% bis-acrylamide gel and subjecting the gel to a low voltage current for 8-16 hours as described in section 2.2.13. This technique was pioneered by Orita et al [208] but has been adapted to allow the detection of single nucleotide polymorphisms as well as microsatellites. Single stranded DNA is affected by factors other than size during its migration through an electrophoretic gel. Single nucleotide polymorphisms (SNPs) within the sequence may also affect the electrostatic nature and therefore potentially the migration of the fragment. Image taken from http://instruct.uwo.ca/biology/366b/SSCPgel.

Markers designed around dinucleotide repeats were prefixed "DNR", those containing trinucleotide repeats were prefixed "TNR", and those surrounding tetranucleotide repeats were prefixed "TetNR".

The microsatellites tested were a rich source of informative markers for the strains involved in the lethal recessive screen. Of the 117 markers which were successfully amplified and tested, 64 showed some level of polymorphism between the strains. There was however, a range of levels of polymorphism between the different strains as can be seen in table 3.3.

	C3H	101/H	B6	BALB/c	sa
C3H		53	33	8	9
101/H	53		50	51	52
B6	33	50		32	31
BALB/c	8	51	32		8
sa	9	52	31	8	

Table 3.3: Polymorphisms identified during the microsatellite screen
The number of polymorphisms found in novel microsatellite markers from the Del36H region within those strains that were used for the lethal recessive screen.

There were distinctive regional patterns of polymorphism across the deleted region (Table 3.4) – some BACs contain microsatellites which were primarily non-informative between the screen strains (e.g.: 80c18, 461o24, 128m23). BACs at the distal and proximal ends of the region contain microsatellite markers which predominantly were not differentiable between C3H, BALB/c and *Satin* strains, but were informative between any one of these strains and B6 (i.e.: 38e20 at the proximal end, and 412p21, 408h17, 235j20 and 262j21 at the distal end. The BAC 41l21 which was also located towards the distal portion of the region showed uniformity in microsatellite sizes between the strains C3H and BALB/c with a different sized product in B6 and 101/H, but the *Satin* strain was not tested with markers derived from the sequence of this BAC).

Strain 101/H exhibited the largest number of microsatellite markers with unique product sizes (a total of 41). However, in markers where 101/H could be differentiated from some strains but not others, 101/H shared its SSCP profile with B6 in 8 of the 10 instances. Seven of these are from the three most distal BACs to be

characterised in the interval. The exception was marker DNR138 in which 101/H is the same size as 3H1 whereas B6 has a unique product size.

SSCP profiles of the mouse species *Mus spretus* and the sub-species *Mus musculus castaneus* were also compared with those of the lethal screen strains. This was because these less related strains were considered for potential future outcrosses.

66 markers were tested for polymorphisms between *M. m. castaneus* and the screen strains. Of these, 9 markers failed to amplify in *M. m. castaneus* and 2 were not informative (DNR006 and DNR009). 74 markers were tested for polymorphisms between *M. spretus* and the screen strains (including all but one of those tested on *M. m. castaneus*), of which 4 failed to amplify in *M. spretus* and 1 was not polymorphic with the screen strains (DNR114 which was not tested in *M. m. castaneus*).

Evidence of haplotype blocks can be seen in these data (table 3.4). These are regions in which closely linked markers exhibit similar or identical patterns of polymorphism between the strains. Tim Wiltshire and colleagues carried out a study of the haplotype patterns seen with SNPs in 8 different laboratory mouse strains and noted that these blocks often extend for tens of megabases and ended abruptly, where they are bordered by recombination hotspots [226]. The polymorphism patterns that can be seen in table 3.4 frequently produced identical haplotypes for 3-4 markers sourced from the same BAC and occasionally also in the closest BAC

from which microsatellites were selected for screening. BACs were screened on average at intervals of once per 5 BACs across the deleted region which equates to intervals of ~668kb given that the MTP of the 12.7Mb region consists of 95 BACs. Haplotype blocks were however smaller in these data than those described by Wiltshire. If as this preliminary analysis suggests, the haplotype blocks within the Del36H interval are smaller than the genome average block size (3.57Mb according to Wiltshire), then the number of recombination hotspots are likely to be more numerous. This could explain the expanded size that was predicted from genetic map data.

The deletion is also larger when estimated cytogenetically than with the physical map. Chromosomes and chromosomal regions have traditionally been determined by measuring their length at metaphase, then estimating the percentage of the whole genome from the size of all the chromosomes combined at an assumed figure of 3000Mb [227]. Some modifications to this approach have since been suggested [228] due to the faster speeds at which the giemsa positive bands condense during metaphase [229] and also due to a more accurate figure for whole genome size since the completion of the mouse genome. Furey and Haussler analysed data from ~9500 previous Fluorescent In-Situ Hybridisation (FISH) BAC mapping experiments and compared these data with previous cytogenetic size estimations. They found that the giemsa positive bands were larger than previous estimations by between 4% and 24%, depending on staining intensity [228].

The deleted cytogenetic bands in the Del36H deletion are from A3.3 to A5 inclusive [21]. Estimations of chromosomal length placed this region at 20.1 % of the length of chromosome 13 and (using a whole genome size estimation of 3.2Gb) – a physical size of 25.6Mb [21]. These figures equate closely with the deletion size predicted by the genetic map but not with the 12.7Mb from the BAC contig sequence.

The structural factors that Furey and Haussler highlighted as sources of chromosome density variability are likely to have contributed to inconsistencies in the estimation of the deletion's length - GC content, repeat structures, CpG island density, gene density and degree of condensation are all likely to effect the apparent size of a chromosomal region during cytogenetic analysis. The deleted region is comparatively rich in G positive band material – of the 3 bands that make up the deleted region, two are G positive (13A3.3 and 13A5) and these are both visually larger than the 13A4 G- negative band (http://www.ensembl.org/Mus_musculus/index.html). This suggests that if Furey's hypothesis is correct, the deletion should be larger than initially estimated whereas BAC sequencing indicated a smaller size.

False Positive Results

Some indication of the level of incorrect scoring from the SSCP and agarose polymorphism screen can be ascertained by comparison of this data with the allele size data determined for a selection of these markers as described in section 3.12.

SSCP occasionally resulted in bands that could be seen in some samples but were hidden in others due to slight differences in signal strength.

BAC RPCI-23 Clone No.	BAC Position in Contig	Total No. Markers Tested	No. of Polymorphic Markers	B6	101	3H1	satin	C3H	BALB/c
38E20	5	4	3	U	U		C ;L	S ;L	C ;S
480B19	21	11	4	U	S;C;L		1;C;L	1;S;L	1;S;C
			1	U	L		U	U	1
			1	1;C;L	B;C;L		U	B;1;L	B;1;C
80c18	26	4	0	N	N	N	N	N	N
461O24	29	4	0	N	N	N	N	N	N
211M15	32	8	1	U			U		
			2	S;C;L	U		B;C;L	B;S;L	B;S;C
19g04	33	3	1	S;C;L	U		B;C;L	B;S;L	B;S;C
169M24	41	2	1	S;C;L	U		B;C;L	B;S;L	B;S;C
189A16	46	3	1	U	U		L	U	S
128M23	54	10	1	U	3;S;C;L	1;S;C;L	1;3;C;L	1;3;S;L	1;3;S;C
262C06	60	12	1	U	3;S;C;L	1;S;C;L	1;3;C;L	1;3;S;L	1;3;S;C
			1	1	B	U	C;L	S;L	S;C
			3	S;C;L	U	U	B;C;L	B;S;L	B;S;C
			1	1;3;S;C	B;3;S;C	B;1;S;C	B;1;3;C	B;1;3;S	U
202F03	63	5	4	S;C;L	U		B;C;L	B;S;L	B;S;C
47A02	65	8	4	S;C;L	U	U	B;C;L	B;S;L	B;S;C
			1	3;S;C	U	B;S;C	B;3;C	B;3;S	U
			1	1;S;L	B;S;L	U	B;1;L	U	B;1;S
235O1	66	8	3	S;C;L	U	U	B;C;L	B;S;L	B;S;C
			1	1;S;L	B;S;L	U	B;1;L	U	B;1;S
232M10	67	9	1	L	U	U	U	U	B
			1	S	U		B	L	C
			3	3;S;C;L	U	B;S;C;L	B;3;C;L	B;3;S;L	B;3;S;C
			2	S;C;L	U	U	B;C;L	B;S;L	B;S;C
412P21	73	3	1	U	U	U	C ;L	S ;L	C ;S
			1	U	3	1	C ;L	S ;L	C ;S
408H17	75	3	3	U	U	U	C ;L	S ;L	C ;S
235J20	76	1	1	U	U	U	C ;L	S ;L	C ;S
322J11	78	8	1	U	U		U	U	U
			1	U	U	U	C ;L	S ;L	C ;S
			2	S;C;L	U	U	B;C;L	B;S;L	B;S;C
			1	S;C;L	U		B;C;L	B;S;L	B;S;C
43L14	86	3	1	U	U	S;C;L	3;C;L	3;S;L	3;S;C
			1	1	B		C ;L	S ;L	C ;S
41L21	93	6	3	1	B			L	C
262J21	95	3	3	1	B		C ;L	S ;L	C ;S

Key to Polymorphism Results

B	Product size = c57BL/6J
1	Product size = 101/H
3	Product size = 3H1
S	Product size = Satin strain
C	Product size = C3H/HeH
L	Product size = BALB/c

	= 1 other product size
U	Unique Product size
N	Not informative
	Not tested or failed to amplify

Table 3.4: Polymorphism data for informative BACs
The number of markers tested are shown by clone, alongside the number of polymorphic markers with each pattern of polymorphism shown. BAC clones are in their respective order within the deletion interval. For complete dataset, see appendix 3.3.

3.8 *Satin* and 3H1 Unexpected SSCP Profiles

3H1 mice are the offspring of matings between C3H and 101/H mice. The SSCP profiles were scored on the basis of whether a difference was seen between profiles, so an individual which was heterozygous for a particular polymorphic allele would be classed as different from either of the two mono-allelic forms. The SSCP profiles for 3H1 were therefore often classified as "unique" when the C3H profile differed from that of 101/H and when the 3H1 strain carried both alleles.

Of the 64 markers tested with 3H1 DNA, 24 (37.5%) showed what was classified as a "unique" product profile. All of these markers were polymorphic between the C3H and 101/H alleles. There were 5 markers which were not classified as "unique" in 3H1 despite a difference between the products of C3H and 101/H. All of these markers shared a profile with C3H rather than 101/H when tested with C3H mice. It is possible that preferential amplification of the C3H allele may have occurred, although it would be surprising if all 5 of the markers preferentially amplified the C3H allele rather than the 101/H allele. Another possibility is that the unique profile of 101/H did not represent a real difference, but again – it is unlikely that the C3H allele would dominate in all 5 of the alleles by such a chance event. These spurious results are unlikely to represent a double recombination because for each example of markers which showed a difference in profile between C3H and 101/H but not C3H and 3H1, there were other markers from the same BAC sequence which did show a difference between C3H and 3H1.

There was a higher than expected variability between 3H1 and *Satin* mice. Of the 37 polymorphic markers which were tested for 3H1 alleles, only 8 produced the same SSCP profile as the *Satin* mice (the same markers for which 3H1 shared a profile pattern with C3H). Fifty of the 58 polymorphic markers that were tested in both *Satin* and C3H produced the same profile, so it appears that the background of the *Satin* strain has more in common with C3H than with 3H1.

The *satin* mutation has a complex history following its discovery by LB Russell in 1955 ([230] referred to in [220]) - prior to the lethal recessive screen it was on a lineage which also contained pearl, muted and vestigial tail. Before and after its re-derivation, *satin* was maintained as a closed colony of *sa/sa* mice segregating C3H and 101/H alleles. However during the screen, the satin strain was bred onto a C3H lineage until it was congenic for C3H. This background may go some way to explain why some regions of the *Satin* genome do not have a 1:1 ratio of C3H and 101/H backgrounds.

3.9 Marker Redundancy

Two markers showed redundancy with other markers in the del36 region: DNR 50 and DNR51 (See appendix 3.2). These microsatellites have almost identical primer sequences, only one base differs between them (the 5th base in the reverse primer is an A in DNR050 and a T in DNR051). It was later discovered that these microsatellites were contained within two related genes which formed a prolactin cluster – DNR50 in prolactin3d3 (Prl3d3) and DNR51 in prolactin 3d1 (Prl3d1).

e-PCR

On screening of the microsatellite sequences with e-PCR (http://www.ncbi.nlm.nih.gov/sutils/e-pcr/), seven of the microsatellites were found to be markers with previous designations.

Redundant Marker	Designated Marker
DNR008	d13mit154
DNR012	d13 Mit 79
DNR023	d13 cwr 7
DNR050	AI325057
DNR051	AI325057
DNR114	D13Mit81
DNR131	PMC315447P1

Table 3.5: Redundant Markers
Markers identified during microsatellite screening which were already designated markers. DNR050 and DNR051 were highlighted as redundant to the same existing marker – AI 325057. This is also reflected in the UNISTS description of AI325057 which warns of matches to multiple loci (http://www.ncbi.nlm.nih.gov/genome/sts/sts.cgi?uid=185969).

3.10 SNP Discovery

SNPs have the advantage over microsatellites that they can be run on systems such as the pyrosequencing unit described in section 2.2.26. It was therefore decided that a pilot study could be carried out to investigate the potential for SNP discovery within the Del36H interval.

Regions of repeat masked BAC sequence were screened with Pipmaker (section 2.2.14) prior to testing, so that only regions of low homology between mouse and human were tested. It was considered important to avoid regions of high homology because this can imply functional significance of a region and therefore reduce the probability of polymorphisms between the strains.

Single base changes can be visualised by SSCP when the denatured (upper) bands are clearly visible. Detection of these bands is more sensitive to differences in salt, temperature, solvents, and ionic strength of the buffer than the lower, dsDNA bands. Because of this, it was decided to run the PCR products in four different conditions. Temperature and buffer type were both varied. Straight TBE was used as well as TBE containing HEPES. An air cooled SSCP tank allowed 20°C runs in addition to the usual runs at 4°C

SSCP gels were run under the various ionic and thermal conditions described in section 2.2.14. This method was used in order to maximise the probability of SNP detection with an SSCP approach.

Primer set c13SNP 1c (c13SNP1 R & c13SNP1c L) produced SSCP bands which suggested a conformational change between B6 and Bc/C3H. This was detected with HEPES/TBE at 4oC and 20oC, and with TBE at 20oC although no difference was observed in TBE at 4oC following standard SSCP running conditions. No polymorphisms were detected in this PCR product.

SSCP of primer set c13SNP 9b (c13SNP9 L & c13SNP9b R) suggested a conformational change between 101/H and all other strains. This was detected at 4oC in both TBE with HEPES and TBE alone. A series of 10 polymorphisms were discovered within this sequence as listed in table 3.6.

Polymorphism Number	Base Change (B6/Bc/C3H to 101/H)	Ensembl v46 Position
1	C-G	29321009
2	C-A	29321083
3	C-T	29321133
4	G-A	29321150
5	G-A	29321155
6	T-C	29321191
7	G-T	29321193
8	T-G	29321199
9	G-T	29321200
10	T-G	29321204

Table 3.6: Single Nucleotide Polymorphisms (SNPs)
These SNPs were discovered in the PCR product c13SNP9b. Ensembl positions are within chromosome 13 and are according to Ensembl version 46.

No SNPs were detected with SSCP in any of the other PCR products. Although several polymorphisms were discovered with this approach, all of them were between 101/H and the other strains of the screen. There were already a large number of microsatellites in the interval with this haplotype so these SNPs did not

provide an additional level of informativeness. The microsatellite polymorphism discovery process was more efficient at the provision of informative markers than was the SNP discovery process – there was no way of telling which regions were likely to contain variability without screening large quantities of PCR products. Conversely, microsatellites are inherently unstable and therefore prone to variation. The relatively small scale of the polymorphism detection aspect of the project meant that SNP discovery was discontinued to allow for a focus on the more efficient microsatellite polymorphism discovery approach. An alternative technology that wasn't available for SNP discovery at the time is dHPLC (described in chapter 4). This would have provided a high throughput and reliable tool for the detection of SNPs.

3.11 Construction of the Genotyping Panels

Marker Panel 1

A total of 17 polymorphic markers were selected from across the region to identify recombination breakpoints (figure 3.4). Seven of the markers were from the Whitehead/MIT marker resource and had previously been tested for polymorphisms by M. Cadman and T. Marsland (unpublished data), 11 were uncovered during the microsatellite polymorphism screen (sections 2.2.14 and 3.7). Allelic sizes of an initial genotyping panel of 10 markers were determined by J. Ragoussis at the Wellcome Trust Centre for Human Genetics (WTCHG), then reanalysed at Harwell. There were some differences between the datasets as can be seen on table 3.7. Some of the size calculation variations could be explained by the use of different capillary electrophoresis systems (the Ragoussis group used the MegaBACE system (Amersham Biosciences), Harwell genotyping was performed with an AB3100 Prism system) and different size standards (the Ragoussis group utilised the ET ROX 400 size standard system (Amersham Biosciences) whereas at Harwell, a NED labelled size standard was used as described in section 2.2.15), but there were some discrepancies which cannot be explained simply on this basis alone.

The same pattern of polymorphism and order of allele sizes was observed in each dataset for the majority of markers. This suggests that most of the variation could result from a difference in the method of size determination by the two systems rather than the amplification of different regions.

DNR003, DNR059 and DNR 024 consist of alternative reverse primers in Harwell set 1 (see table 3.8) whereas the original reverse sequence was used at the WTCHG. This explains the large differences between the product sizes seen in DNR003 and DNR059 at the two centres. The marker TetNR10 showed a size difference of 9bp between the Harwell and WTCHG sets. This difference was exactly the same in each strain – and the Ensembl size prediction for the B6 allele (212bp) was closer to the Harwell prediction of 211bp than the WTCHG prediction of 220, suggesting that different primers may have been used at the WTCHG to amplify the same microsatellite. It is also possible however, that this difference is due to the differences in systems and size standards.

D13Mit17 has a difference of 8.2 bp for the B6 allele, but less than 4 for the others. The Harwell B6 prediction of 169bp appears closer to the Ensembl prediction of 168bp than does the WTCHG prediction of 177.2. The differences seen in these size calculations are small enough that they could be due to size standard or system differences.

The initial Harwell panel was replaced by a second panel of markers after the third plate of 96 samples were genotyped. The primary reason for this was because there was a large gap (8.25Mb) between the markers DNR 59 and TetNR5. In order to achieve an improved breakpoint resolution in the event of recombination and to reduce the risk of double recombinants going undetected, a number of additional markers were inserted into this gap.

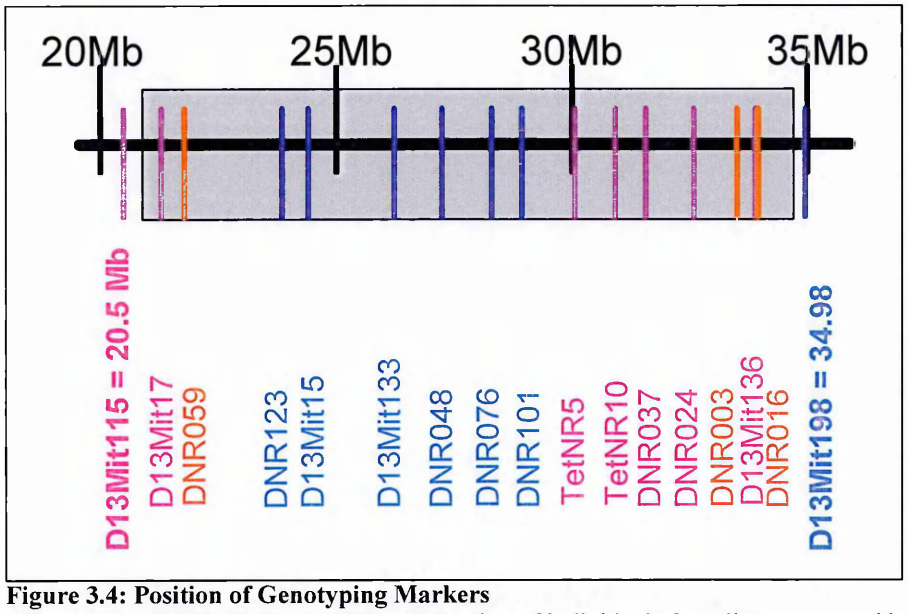


Figure 3.4: Position of Genotyping Markers
The position of each marker used for genotyping of individuals from lines generated by the lethal recessive screen for mutations within the Del36H region. These positions were calculated by Ensembl version 41 as shown in appendix 3.2. Red markers are those markers which were in the first panel but removed from the second panel. Pink markers were in both the first and second panels. Blue markers were only present in the second panel. The Black divisions represent intervals of 5Mb across the region.

Del36 Genotyping Marker Sizes (bp) and Dyes

Marker Panel	Mouse Strain	D13Mit115 FAM	D13Mit17 FAM	DNR059* ROX	DNR123* FAM	D13Mit15* FAM	D13Mit133 HEX	DNR48* HEX	DNR76 FAM	DNR101* HEX	TetNR5 FAM	TetNR10 FAM	DNR37 ROX	DNR24* HEX	DNR003* HEX	DNR016 ROX	D13Mit136 FAM	D13Mit198 ROX
Wellcome Panel	B6	147.51	177.20	209.71							311.60	220.00	223.62	193.80	214.38	170.23	92.79	
	C3H	138.09	154.95	205.46							311.00	216.00	233.54	187.97	218.54	188.06	94.70	
	101	141.06	168.69	201.45							299.00	220.00	231.00	207.00	214.31	170.47	92.70	
	BALB/c	133.89	154.95	205.44							311.80	216.00	229.66	188.02	218.52	188.06	94.70	
Harwell Panel 1	B6	147	169	337							308	211	219	181	289	178	96	
	C3H	138	153	332							316	207	229	173	293	197	98	
	101	140	165	328							296	211	225	195	289	178	96	
	BALB/c	134	153	332							308	207	225	173	293	197	98	
	sa	136	153	332							319	207	221	173	293	197	98	
Harwell Panel 2	B6	146	169		276	234	152	264	182	211	308	212	220	181			96	138
	C3H	138	153		299	227	148	273	186	207	317	207	229	173			98	140
	101	140	165		296	227	156	267	186	207	296	212	225	195			96	138
	BALB/c	134	153		296	237	148	272	186	207	308	207	225	173			98	140
	Sa	136	153		299	227	148	272	186	207	320	207	222	173			98	140

Table 3.7: Marker Panels Used in the different genotyping schemes
Markers used to genotype individuals arising from the lethal screen. Underlined markers are those which were modified to produce different product sizes for the Harwell panels as described in table 3.8. Each forward primer was labelled with a fluorescent dye as indicated by the highlighted colour and by the dye listed below each marker heading. The figures from Harwell are medians of the size ranges seen for each marker (all markers had a range of less than +/- 1.)

Marker Panel 2

The markers DNR 059, DNR003 and DNR 016 were removed from the second panel because adjacent markers were available which had the same level of informativeness.

DNR 59, although informative between B6, 101/H and BALB/c/C3H/*Satin*, is only 0.5 Mb from D13Mit17 (figure 3.4) and has the same level of informativeness as this marker (see table 3.7). DNR003 and DNR016 are both very close to marker D13Mit136 (0.41Mb and 0.002Mb respectively) and all three had the same level of informativeness (101/H $\neq$ B6 $\neq$ C3H = Bc = *Satin*). In order to allow an easy transition to genotyping with the increased number of markers, DNR059, DNR003 and DNR016 were removed from panel 2.

It was intended that D13Mit117 should be added to the second panel of markers. This marker (at 34980274bp Ensembl v.41) lies outside the deletion at the distal end (the deletion breakpoints were determined by the Ragoussis group to 20870225-34915188 in Ensembl v41 (20954629 – 34999592 in Ensembl v50) but is more informative than D13Mit198 because it is polymorphic between B6 and 101/H, in addition to differentiating between those strains and the other screen strains. Unfortunately optimisation with the multiplex PCR conditions used for this aspect of the project was unsuccessful and D13Mit117 was not amplified in the presence of the other HEX labelled primers.

DNR111(Ri) was a variation of marker DNR111 with an alternative reverse primer that produced an appropriately sized product for incorporation into the Harwell marker set (see table 3.8). This was a ROX labelled marker that was intended for multiplex PCR with DNR 37 and D13Mit15 (F+Ri), but it failed to amplify in their presence. Although DNR111 lies close to DNR 101 (0.3Mb separates the markers), its presence in the panel could have increased informativeness because it distinguishes between 101/H and B6. Unfortunately, despite apparently successful optimisation in a preliminary study, the product was not robust in response to small variations in DNA quality so very few individuals were successfully genotyped for this marker.

As can be seen from tables 3.4 and 3.7, there is a varying degree of informativeness across the deleted region. The absence of a panel of fully informative markers limits genotyping confidence when screening individuals from a screen as complex, and with as many strains involved, as this type of screen. It was therefore not possible to specify the genetic origin of each locus.

The mutagenised or potentially mutagenised parental chromosome in each of the mutant lines should be derived from *Satin* - 3H1. The reason that both C3H and 101/H were tested for polymorphism is because 3H1 consists of both C3H and 101/H alleles. Any recombinations between chromosomes derived from these strains is not significant for the purposes of this study because either may have been mutated whereas B6 and BALB/c should be wildtype.

Alleles derived from BALB/c, C3H and *Satin* show similar profiles at several markers. This has meant that where a recombination has occurred between BALB/c and C3H, there cannot be as high a degree of confidence in determining the breakpoint position as there would be in a recombination between B6 and C3H. B6 parental origin could be uniquely determined at 10 of the 14 loci selected for the second panel of polymorphic markers allowing higher degree of accuracy in determining the breakpoints between B6 and mutagenised chromosomes.

Optimising product size of genotyping markers

Four of the panel 1 markers required further optimisation for a robust PCR and clear distance between marker band sizes (on agarose optimisation tests) in multiplex conditions: DNR016, DNR059, DNR003 and DNR024. Alternative reverse primers for these markers are shown in table 3.8.

Due to the increased number of markers in the second panel, there were a number of markers which shared product sizes. B6 product sizes could be predicted from the available sequence data, but the exact sizes of products for each of the other strains were not known.

Reverse primers were therefore designed for a range of markers that overlapped or were of a similar size when amplified. This approach maximised the combination of

different sized products and was intended to allow the construction of a marker panel with clear differences between each product within the dye sets.

PCR optimisation was carried out with the original dye labelled forward primers and each of the reverse primers individually - in order to achieve clear, strong signals for each product. The estimated sizes were used to construct batches of markers for genotyping with discernable distances between each product when run on an agarose gel.

Four markers were altered for genotyping panel 2 (see table 3.8) and one, DNR059 was altered for genotyping panel 1, due to a more robust PCR and a greater difference between marker product sizes with this sequence.

Primer	Alternative Reverse Sequence 1 (Ri)	Alternative Reverse Sequence 2 (Rii)	Alternative Reverse Sequence 3 (Riii)
DNR076	ACAAAGAAGTGCTCCCAACC	CCCAGCTTAGCAAACCAATAC	ACAGGTGTGCATCAACAGAAC
DNR123	CGGCAGGATGCCTATTATATTG	TGATAAGTAGTTCCAGGCCTTAGC	CCACAGATCATCTAGCTTGTGC
DNR101	TTTCTTCCGCTTCCATTAGG	CTTCCAACTGGGCCAAAAC	TCCATTAGGCAGATCAAATGTC
DNR048	GCAGATTCAGGTGTCTGAGG	GGGTTACTTGTATCACCCTACCC	GAGGGTTATTGAGATCTTGATGG
DNR111	AACAGAGACCCTTTGTGAGAGC	GACTCCTGAGTGGTTGTGAGC	AAGCACAGCAGGAAAACTAACAG
DNR016	TCCTAAGCCACACCCAAGAG	ATCCCTGGAACCTCAAATCC	
DNR059	TTTTGCCCTATCCAGACCAG	CGGAAAAGGAAAAGAATAAAAGG	
DNR003	ATAAGGCATTGGCGAGAATC	TAAAGGGCATTTCCTTGGTC	
DNR024	TCACTCAACACTTACATCATAGGC	GGGGTACCTCAATCACTCAAC	
D13Mit15	CCGTTTAGGTTTTTAAGTTACAGG		

Table 3.8: Alternative reverse primers for genotyping primers
The primer sequences in red were those which replaced the original reverse primers for the genotyping screen i.e.: DNR123 Rii, DNR101 Rii, DNR048 Ri, DNR024 Rii, D13Mit15 Ri. DNR111 was not included in the genotyping panels due to the lack of a robust product for PCR with any of the reverse primers.

3.12 Analysis of Genotyping Data

The nature of this genotyping project and the number of strains involved increased the possibilities for ambiguities within the data. There were effectively four strains which required differentiation during the genotyping of individuals from the lethal recessive screen. For simplicity, *satin* alleles were characterised as "C3H". Most of the markers would not differentiate between these two origins and for the purposes of the project the difference was irrelevant because the allele sizes of C3H and *sa* both indicate origin from the mutagenised lineage. The two strains were therefore grouped together for convenience of data organisation.

Ideally, each marker used for characterising the genetic origin of the Del36H region would be polymorphic between each of the in-bred mouse strains involved in the screen. This is unrealistic however given the close degree of relatedness of some of these lineages. It was therefore necessary to select a panel of markers which provided the maximum level of polymorphism between the strains selected from the available set, at approximately regular intervals across the interval.

In order to simplify genotyping output, the combined informativeness of markers was employed in order to regionally determine the specific strain origin across the deletion. In recombinant individuals, strain origin was frequently ambiguous but strain information could be inferred by utilising the differing polymorphisms of the surrounding markers, thus limiting the region of ambiguity.

For example, the haplotype pattern shown in the genotyping data of figure3.5A shows a (C3H x B6)F1 individual. Five of the 10 markers in panel 1 (and 10 of 14 markers in panel 2) can differentiate B6 from the other strains and it is clear that B6 is present across the region on one of the two chromosomes. Many of the markers though, cannot differentiate between C3H and BALB/c, but those that do differentiate suggest the presence of C3H and not BALB/c therefore this individual was scored as a non recombinant individual.

		d13mit115	d13mit17	DNR059	TETNR05	TETNR10	DNR037	DNR024	DNR003	DNR016	d13mit136
A	Potential Strains 1	B	B	B	B/L	B/1	B	B	B/1	B/1	B/1
	Assumed Strain 1	B	B	B	B	B	B	B	B	B	B
	Allele Size 1	148	177	210	312	220	224	194	214	170	93
	Allele Size 2	138	155	205	255	216	234	188	219	188	95
	Assumed Strain 2	C	C	C	C	C	C	C	C	C	C
	Potential Strains 2	C	L/C	L/C	C	C/L	C	C/L	C/L	C/L	C/L
B	Potential Strains 1	B	B	B	B/L	B/1	B	B	B/1	B/1	B/1
	Assumed Strain 1	B	B	B	B	B	B	B	B	B	B
	Allele Size 1	148	177	210	312	220	224	194	214	170	93
	Allele Size 2	134	155	205	255	216	234	188	219	188	95
	Assumed Strain 2	L	L/C	L/ C	C	C	C	C	C	C	C
	Potential Strains 2	L	L/C	L/C	C	L/C	C	L/C	L/C	L/C	L/C

Figure 3.5: Example of Del36H Genotype Determination
Marker combinations were used to infer strain origin within regions of the interval. A is a non-recombinant B6/C3H individual (D36/1409.01 in appendix 3.1). B is a recombinant individual with B6 and a recombination between BALB/c and C3H. (Mud/1616.05 in appendix 3.1). Note that double recombinants between these strains could be missed due to a lack of markers that can be used to distinguish between BALB/c and C3H.

Although a large number of markers were screened, there remains a dearth of markers which are informative between BALB/c and C3H. DNR037 and TetNr05 were uncovered during the screen, and there are several markers that can

differentiate BALB/c from either *satin* alleles or C3H alleles but not both. However, this was not of interest within the context of this project because genotyping was performed in order to identify the alleles on chromosomal segments which could potentially carry the mutation (C3H, *Satin*, 101/H) from those which wouldn't (B6 and BALB/c) in order to define the critical regions in which mutations were most likely to be found.

Ambiguous 296bp FAM Product

An oversight in the design of marker panel 2 resulted in a FAM product at 296bp which could be either a DNR123 product of 101/H origin or a TetNR05 product of BALB/c or 101/H origin. This was a potential problem for unambiguous genotyping. In hindsight, one of the primers for these markers should have been reordered with a different dye (either ROX or FAM) label in order to benefit from the maximum informativeness of the panel.

Fortunately, the markers on either side of DNR123 (D13Mit15 and D13Mit17) were informative for BALB/c and 101/H when combined. The combination of markers adjacent to TetNR05 (TetNR10 and DNR101) was also informative for 101/H. In addition to this, many of the other markers within the panel are also informative for 101/H.

Of the 741 suspected recombinants from genotyping at all stages, only 25 individuals carried a BALB/c or 101/H allele at DNR123 or carried a 101/H allele at

TetNR05. Most of the BALB/c and/or 101/H alleles at these loci fitted the general pattern of the region and could be confirmed as either both loci of 101/H origin, both of BALB/c origin, or one loci of BALB/c and one of 101/H origin – by analysis of the adjacent markers.

There were three individuals, however, for which the information from the surrounding markers did not resolve the ambiguity at these loci.

The two individuals with the genotype seen in M1239.29 (the 29th listed haplotype for animals of line 1239 after its confirmation as a heritable mutation) of appendix 3.1 were scored as 101/H at DNR123 and C3H at TetNR05. However, the size overlap of the FAM product at 296bp would suggest that a 101/H allele could also be present at the TetNR05 locus. This ambiguity would not affect the significance of the result as either 101/H or C3H could represent the mutagenised background of 3H1.

The one individual with the haplotype indicated by M412.22 in appendix 3.1 was scored as B6 at DNR123, which is the most distal marker of a region of B6 (adjacent to a region of B6/101/H), and B6/101/H at TetNR05. The presence of the 101/H allele at TetNR05 suggests that a 101/H allele could also be present at DNR123. However the marker proximal of DNR123 is D13Mit17 – this was conclusively B6 without the presence of a 101/H allele so a region of only B6 origin is definitely present in this individual but the extent of this region is uncertain. In this case the

size ambiguity limited the resolution of the breakpoint position following a recombination between B6 and 101/H.

One event which could have been missed as a result of the overlapping product sizes is a double recombination leading to 101/H or BALB/c at DNR 123, or 101/H at TetNR05, between the surrounding markers. However, given the relatively low numbers of individual genotyped with Harwell marker panel 2 that carried either BALB/c or 101/H alleles within the interval, this seems an unlikely occurrence.

Genotyping

Genotyping of each of the mutant lines was performed in collaboration with J. Ragoussis and analysis of the combined results was performed by D. Bogani and C. Willoughby. All animal work in this section was performed by the Early Development group and animal staff from building 383 (MRC Harwell). Data from 1325 animals genotyped at MRC Harwell was added to data from 6446 animals genotyped by the J. Ragoussis laboratory for a total of 7771 animals. The screening of mutant lines was carried out in three different stages: Lethality testing, inheritance testing, and mapping.

Lethality testing

Animals were bred for lethality testing as described in figure 3.2. Genotyping was carried out to ensure that the mutagenised chromosome wasn't disrupted by recombination within the Del36H interval in animals to be used for the

determination of embryonic lethality. The exact one-tailed binomial test was used to compare the birth-rate of *satin* coated animals with the expected birth-rate. This allowed the determination of whether a lethal recessive mutation was present in the lineage or not (D. Bogani, personal communication). Individuals with the prefix "Mud" in appendix 3.1 were those animals resulting from the lethality testing stage of screening.

Inheritance testing

The inheritance testing stage also required genotyping to ensure that the mutation wasn't lost due to recombination. *Satin* coat locus data from all of the individuals of each cross was subjected to the one tail binomial test at 95% and 99% significance levels. It was at this stage that the severity and dominance relationship of the mutation could be determined. Individuals with the prefix "Mudy" in appendix 3.1 were those animals resulting from the inheritance testing stage of screening.

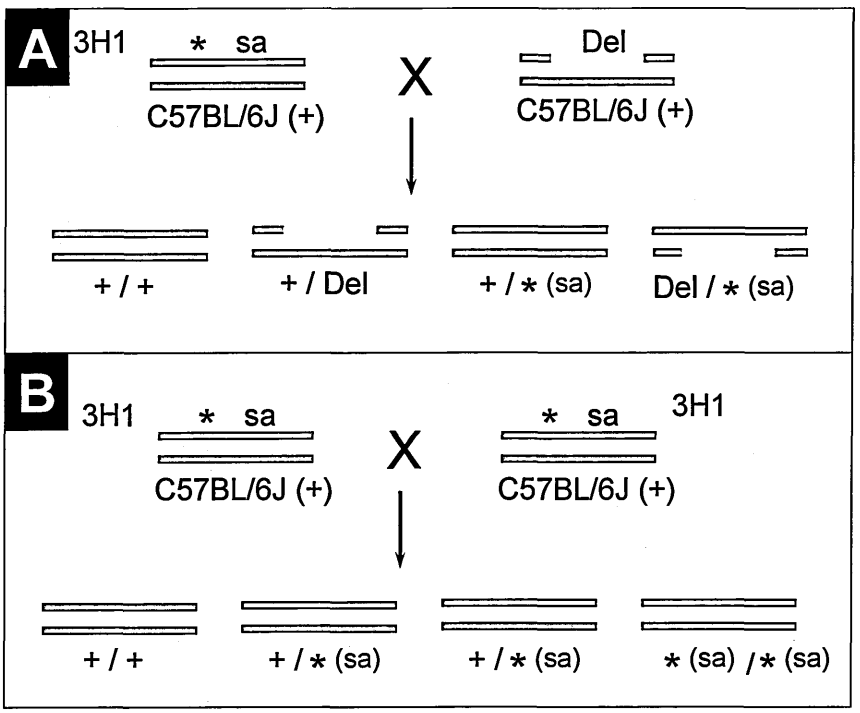


Figure 3.6: Breeding Schemes for Inheritance Testing
Two breeding schemes were used for the inheritance testing stage of the lethal recessive screen. *+/sa* heterozygotes were either intercrossed or crossed to Del36H deletion carriers until at least 60 progeny were born and tested for the *Satin* coat locus from each line.

Mapping and Mutant Line Maintenance

Mutants were crossed with Del36H deletion carriers and offspring were genotyped to identify recombinants. Recombinant individuals were mated as illustrated in figure 3.6 and the numbers of *Satin* progeny from recombinant lines underwent the exact one-tailed binomial test to determine if the mutation effect persisted. This approach enabled mutations with varying levels of penetrance to be mapped.

These offspring, in addition to animals which were bred for the maintenance of proven mutant lines were designated with the prefix "m". Genotyping was required to ensure that haplotypes were maintained between generations during recombinant

mapping and colony maintenance. The majority of animals with this designation listed in appendix 3.1 were genotyped for colony maintenance.

The mutations in the mutant lines were localised to the following regions (as shown in figure 3.7):

54: Between TetNR10 and D13Mit 136.

91: Sox4.

241: Between D13Mit17 and 15.

369: Between TetNR10 and D13Mit136.

412: Between TetNR5 and D13Mit136.

624: Between TetNR5 and DNR24.

876: Between DNR24 and D13Mit136.

1073: Between D13Mit133 and D13Mit136.

1185: Between TetNR10 and D13Mit 136.

1616: Between D13Mit84 and D13Mit136.

1239: Between D13Mit17 and 15.

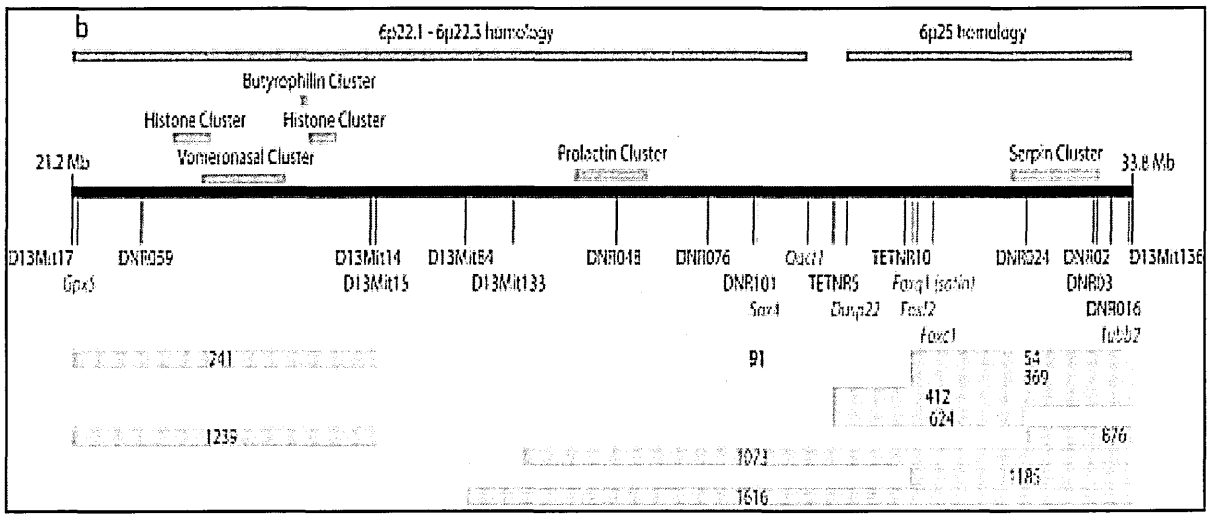


Figure 3.7: Critical Regions of the Del 36H Mutants
Critical regions of the 11 mutant lines identified during the lethal recessive screen, including the line 91 mutation which was mapped to Sox4. Mapping analysis of each of the mutant lines except line 91was performed by C. Willoughby and D. Bogani using the genotyping data shown in appendix 3.1 combined with genotyping data from J. Ragoussis. Image taken from Bogani et al [26].

3.13 Line 91 Recombinant Genotyping

Line 91 homozygous mutant embryos died at 14.5 dpc due to circulation defects caused by malformation of the atria and ventricular septa, a dysplastic mitral valve, and a reduced aortic arch [26]. The mutation was mapped to Sox4 (Ser70Pro) by M. Cadman [26].

Three surviving mud91 recombinant B6 / 3H1 deletion carriers were identified by J. Davies with the genotyping markers D13Mit115, D13Mit133 and D13Mit 17. Further analysis of these individuals with the novel SSLP markers identified in section 3.7 indicated that the mud91 mutation lay between TetNR5 and DNR 201 (see table 3.9). This is a physical distance of approximately 4.97Mb according to the Ensembl marker positions as listed in appendix 3.2. Frequency of recombination placed the mutation within 0.8cM of *Foxq1*, further defining the critical region [26].

Each of the recombinants that were mapped survived, so the region containing the mutation which would have occurred on a *Satin* chromosome has been replaced with B6 derived alleles in these individuals.

Marker	BAC	R1	R2	R3	BAC Order
DNR59,63	38E20	B6			5
DNR60	38E20	B6	*sa*	*sa*	5
DNR123	480B19	B6	*sa*	*sa*	21
DNR201	211M15	B6	B6	*sa*	32
D13Mit133	201H16	B6	B6	B6	39
DNR48	189A16		B6	B6	46
DNR76	128M23	B6			54
DNR101	262C06	B6			60
DNR111	262C06	B6			60
TetNR5	232M10	*sa*	B6	B6	67
DNR138	412P21	*sa*			73
TetNR10	412P21	*sa*			73
DNR152,157,158	408H17	*sa*			75
DNR173	235J20	*sa*			76
DNR37	322J11	*sa*	B6	B6	78
DNR 21,24	43L14	*sa*			86
DNR 2,3,4	41L21	*sa*			93
DNR 16,17	262J21	*sa*			95

Table 3.9: Genotyping of Mud91 for Confirmation of Sox4 Localisation
Mud91 is a mutant line derived from the lethal Recessive screen. R1, R2 and R3 are recombinant individuals which were initially identified by genotyping with the markers D13Mit17 and D13Mit136 which occur near opposite ends of the deleted interval. These data reduced the critical region to between the markers DNR201 and TetNR5 which is an interval of 4.97Mb. This region included the transcription factor Sox4 (on BAC 262C06) in which a mutation had previously been identified in line 91 – the recombinant mapping therefore supported the hypothesis that this mutation was responsible for the line 91 phenotype.

3.14 Foxf2 Mutation Detection

Foxf2 lies within the Del13(Svea)36H deleted region and was considered a strong candidate for several of the recessive lethal mutant lines. Midline development defects have been observed in the form of cleft palate in *Foxf2* knockouts [153]. Although this was not as severe as the phenotypes seen in any of the lines from the lethal screen, holoprosencephaly is caused by the incomplete cleavage of the embryonic forebrain - which is also the result of a midline defect. It was therefore considered that a common mechanism may underlie these mutant forms. The occurrence of holoprosencephaly at varying degrees of severity in lines 54, 241, 369, 624, 1073 and 1239 was considered as evidence for the potential involvement of *Foxf2* in mutations from these lines.

DNA from one confirmed mutation carrier of each of the lethal lines except line 91 was sequenced for mutations in the coding sequence of *Foxf2*. Although genotyping data would suggest the exclusion of *Foxf2* as a potential candidate from several of these lines, the low penetrance of some phenotypes did not allow such exclusions with a high degree of confidence. Mistakes in map location could have occurred with these low numbers, so all of the mutants that weren't already cloned were sequenced. Some of these lines did not show a holoprosencephaly phenotype, but the transcriptional activity of this gene combined with its wide expression throughout the developing embryo suggested that it potentially has important and diverse roles in the regulation of development that have not yet been identified. *Foxf2* maintains a persistent presence within a highly conserved cluster of forkhead box genes throughout the bilaterian group [126], and the benefit of potentially expanding the

allelic series that are described in sections 4.5-4.6 meant that it was worth investigating each of the lines for mutations in this gene. Comparative sequence analysis of each of the different lines was performed on Contig Express before alignment with the consensus mouse genome sequence.

Despite the strong candidacy of the gene, no mutations were identified in the open reading frame of *Foxf2* in any of the lethal lines. The absence of a *Foxf2* coding sequence allele however, does not mean that *Foxf2* is not critical for the survival of early embryos. It has been predicted that with the analysis of 1730 ENU mutagenised pedigrees, the probability of a functional allele being present within any one of the deletion interval genes is 80% [26, 201].

This calculation though, is based on the assumption that a functional allele would be detected. The selective parameters of a phenotype approach such as a lethal recessive screen are much narrower than the genotype-based approach that was used in the calculation of these probabilities – A gene which has the potential for causing a prenatal lethal phenotype if mutated could instead become mutated at a locus which results in a less severe phenotype. So there is less than an 80% chance of a lethal phenotype being discovered in any potentially lethal gene in the region unless all potential coding mutations in each of these genes would result in lethality.

Foxf2 has not yet been thoroughly characterised, and its expression in embryonic tissues such as the CNS, ear, eye and limb buds where its close paralogue Foxf1 is

not expressed, suggest important roles for the gene in the developing embryo. Even within the mutants for which the open reading frame of Foxf2 has been sequenced, it may also be worth sequencing the UTRs and it is certainly worth investigating Foxf2 promoter regions.

Chapter 4: Foxf2 Mutation Detection

4.1 Introduction

Foxf2 was screened with Two ENU mouse mutagenesis archives to increase the probability of obtaining an informative allelic series for the gene. It was considered a strong candidate for the mutations that were envisaged could be uncovered using the deletion-based lethal recessive screening approach. Candidate selection before mutation discovery was possible because the nature of the screen meant that any mutations would be pre-mapped to within the 12.7Mb deleted region, and would result in embryonic lethality. Therefore genes in the deletion interval that were implicated in early development were key potential targets to screen the ENU archives.

Forkhead box (Fox) transcription factors are frequently involved in cell fate determination and pattern formation during organogenesis. Foxf2 is one member of a tightly linked and highly conserved cluster of three Fox genes in the Del36H region on proximal mouse chromosome 13 (consisting of Foxc1, Foxf2 and Foxq1). It shares close sequence homology with Foxf1 which is located in a semi paralogous cluster on mouse chromosome 8 (Foxl1, Foxc2 and Foxf1).

Foxf1 and Foxf2 Expression and Sequence Homology

The DNA binding domain of Foxf1 has 100% sequence identity with the DNA binding domain of Foxf2. Aitola and co-workers carried out a systematic comparison of the expression patterns of the two genes [188]. They found a significant overlap between the expression of the genes during embryonic development in the alimentary, respiratory and urinary tracts. However in contrast to Foxf1, Foxf2 is expressed in the central nervous system (CNS), sensory organs and limb buds; and is absent from the liver capsule. It has also recently been shown to have a role in the regulation of potassium channel formation in heart muscle [231].

Ormestad has shown that Foxf1 is present at high levels throughout the gut whereas Foxf2 is more localised to the oral cavity and posterior gut [232]. He also detected much higher levels of Foxf1 expression in the extraembryonic and lateral plate mesoderm, and suggested that this was a possible explanation for the greater severity of the Foxf1 mutant phenotype compared with the Foxf2 knockout which was described by Wang et al [153].

Foxf2 in the eye

Foxf2 expression was observed in eye tissue in both the developing embryo and adults [188] and was considered a candidate for the gummy eyes and irregular pupil phenotype that was observed in the Del36H hemizygotes [21, 82]. Foxc1 was already known as an important gene in eye development due to the discovery of Foxc1 mutations associated with Axenfeld-Rieger anomaly [130, 131], including

incorrectly aligned corneal cells and hypoplastic stromal mesenchyme and iris pigmented epithelium [81]. However, the eye phenotype observed in the Del36H animal did not match that of known Foxc1 mutants and work carried out by Nishimura and colleagues suggested linkage of a heritable glaucoma (and therefore eye development locus) to an additional gene within human 6p25 [131].

Mutations in a number of different forkhead genes such as human FOXE3, FOXL2 and FOXC2.cause a wide spectrum of different eye disorders in humans[126]. The expression of Foxf2 in the eye [188] and the further localisation to the mesenchyme adjacent to the optic stalks [232] highlighted this gene as an immediate candidate for the further analysis of the eye defect.

The selection of Foxf2 for the gene-driven screen approach was vindicated to some degree by the subsequent discovery of holoprosencephaly in lines 54, 241, 412, 624, 1073 and 1239 from the deletion screen. Holoprosencephaly is caused by mid-line defects and is characterised by the incomplete cleavage of the embryonic forebrain. Palate malformations are also caused by midline defects and these have since been discovered in Foxf2 knockouts [153]. This highlighted Foxf2 as a key candidate for the above mutant lines, although no open reading frame mutations were present in Foxf2 for any of these lineages (as discussed in section 3.14).

4.2 Why Search for an Allelic Series?

The gene driven screen approach was envisaged as having the potential to complement the deletion screen. It would maximise the probability for obtaining an allelic series of Foxf2 – one of the least well understood candidates in the region.

An allelic series of point mutations within a gene can provide a richer source of information about the functional activity of a gene than only a single point mutation or gene knockout. A range of effects on protein function caused by different point mutations within a gene could include amorphic mutations (loss of function), hypomorphic (partial loss of function), antimorphic (opposing/dominant negative function), hypermorphic (exaggerated function), and neomorphic mutations (novel gain of function)[233].

The value of an allelic series in the analysis of the functional activity of a gene was highlighted by Roger Cox and co-workers [234], who localised five mutant alleles to the quaking gene. These alleles caused a range of phenotypes from early embryonic lethality to uncontrolled tremors that were the result of dysmyelination of the CNS. Each of the alleles differed in some way from the others – there was a spectrum of tremor severity between the mutations; the deaths of embryos occurred at differing stages of development; and male infertility resulted from one of the mutations.

The Gene Driven Screen & The ENU Archives

The knockout mouse project [235, 236] aims to induce targeted knockout mutations at every single gene in the mouse genome and to analyse the result that this has on the phenotype.

The Gene driven screen approach requires archives of mutagenised animal lines and a parallel archive of their DNA. The animal lines can take the form of either: sperm [201] [202], embryonic stem cells [237] or live animals [238]. The DNA is screened for mutations, and if any are discovered, animal carriers are available for mutation characterisation. This approach has the advantage over the phenotype driven approach that specific genes of interest can be targeted, and it is possible to proceed directly from sequence information to the isolation of mutations.

A mutagenesis programme was undertaken at MRC Harwell which utilized the mutagen ENU in a phenotypic screen of 26,000 F1 progeny from crosses between C3H females and mutagenised BALB/c males [55].

DNA was extracted from tail biopsies of 2230 of these F1 individuals and archived in conjunction with sperm samples taken from the same individuals. A pilot study demonstrated the feasibility of the high throughput screening of this DNA, followed by the recovery of any identified mutants from the sperm bank [201]. This initial resource has since been expanded by the ongoing construction of an additional MRC

Harwell archive of over 4000 further individuals, greatly increasing the probability of detecting one or more functional allele(s) [202].

4.3 Identification of Mutations with Denaturing HPLC

Two DNA archives have been produced at MRC Harwell. The first was made in collaboration with Glaxo-Smithkline [201] but the second larger archive was produced by the MRC alone [202]. Both of these archives were screened for mutations within the gene Foxf2 using Denaturing High Powered Liquid Chromatography (dHPLC).

Pooled PCR product from the region of interest is denatured then hybridised when using this approach. Opposing strands from each individual will anneal. If however, there is a sequence variant or SNP in one individual, some of the opposing strands in the pooled sample would contain non-complementary base-pairs (heteroduplexes) which would repulse each other and weaken the bond between the two strands as shown in figure 4.1.

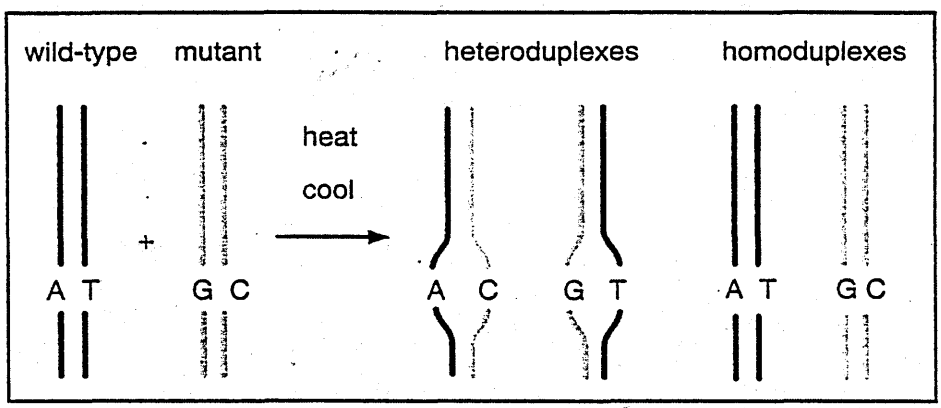


Figure 4.1: Heteroduplexes in dHPLC Mutation Detection
Heteroduplexes cause the separation of DNA strands at the site of a mutation and therefore the two strands have a reduced affinity for each other.

After hybridisation, the pooled PCR product is injected into the dHPLC where it forms a complex with the TEA+ ions of TEAA (triethylammonium acetate) that binds to the dHPLC column.

As the run progresses, the buffers which are pumped through the column, contain a gradually increasing percentage of acetonitrile. This eventually disrupts the chemical interaction between the TEAA-DNA complex and the column. Heteroduplexes elute from the column first when the column temperature is high enough to partially denature the region of the fragment that contains the sequence variation.

One potential problem which can occur during the screening of F1 archives such as those produced by MRC Harwell and Glaxo-SmithKline, is that sequence polymorphisms between the strains may interfere with the screening process. Heteroduplexes would form on fragments which contain these polymorphisms, resulting in additional peaks on a dHPLC profile. However, once a polymorphism is

identified as such, the peak profile can still be compared with the profiles of later samples. Despite the reduced detection efficiency, fragments which contain such polymorphisms can still reveal the presence of point mutations (M. Quwailid, personal communication). .

4.4 Optimum Temperature Conditions for Mutation Detection

The optimum temperature conditions for mutation detection recommended by Transgenomic require that the helical fractions are between 70%-85% [239]. As can be seen in figure 4.2, these specifications mean that a very small percentage of the fragment is at the optimum melting temperature for mutation detection at any one temperature.

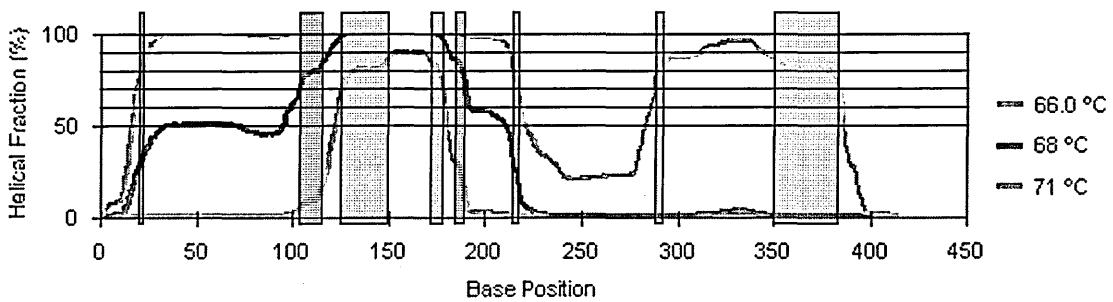


Figure 4.2: Optimal Detection Limits of the Foxf2_1_1 Fragment
Regions within the Foxf2_1_1 fragment which would have the recommended optimum helical fractions of between 70-85% at 66°C, 68°C and 71°C. Highlighted regions represent regions which have a helical fraction of between 70-85% - Pink regions: at 71°C, blue regions: at 68°C, red regions: at 66°C

However, Transgenomic suggested that in practice, less stringent criteria for temperature selection of a helical fraction of between 60% to just under 100% would be acceptable. These less stringent conditions mean that a significantly larger

fraction of each fragment could be considered to have a good chance of having any mutations within them detected.

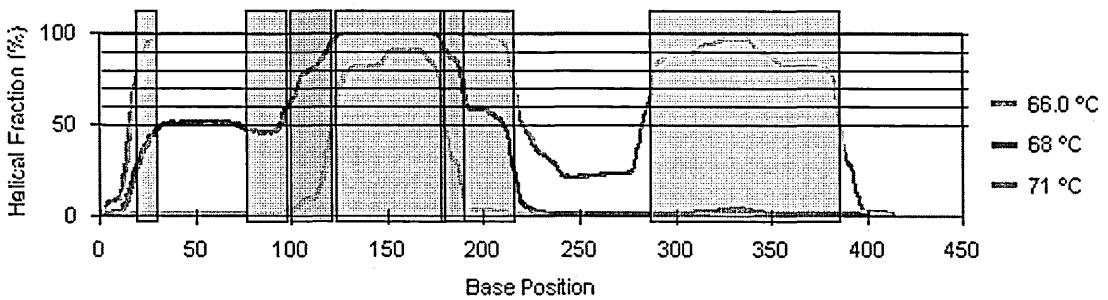


Figure 4.3: Relaxation of dHPLC Helical Fraction Limits
Regions within the Foxf2_1_1 fragment which would have helical fractions of between 60-99% at 66°C, 68°C and 71°C, as suggested by Transgenomic for a high probability of mutation detection.

These specifications would still require a large number of temperature settings in order to gain full coverage of each fragment, this would prove costly and time consuming. Even under these more lenient conditions, many fragments would contain regions where it was impossible to achieve such helical fractions using the WAVE machines in which temperatures can only be altered in increments of 1°C. However, data from Transgenomic suggests that mutations can be detected at a wide range of temperatures outside optimum denaturing temperature (1°C higher and 5°C lower than the optimum temperature) (see figure 4.4). This implies that a 1°C leeway could be allowed without reducing mutation detection probability by any significant degree.

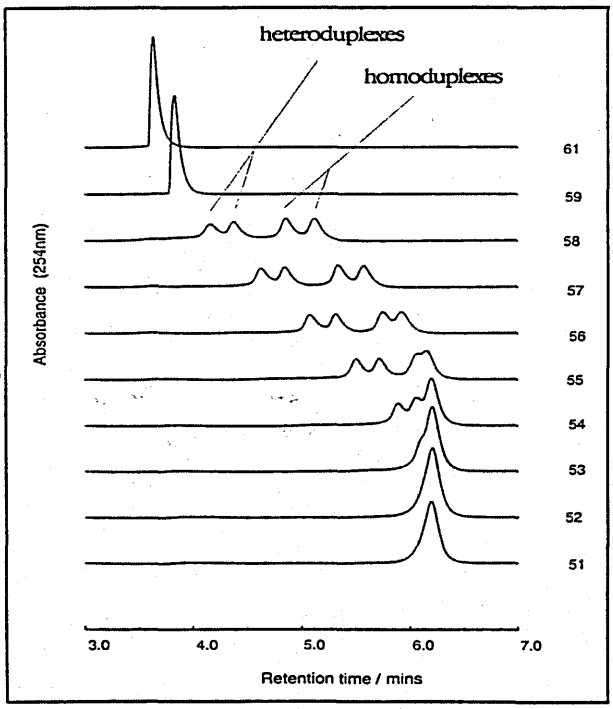


Figure 4.4 Temperature affects mutation detection
Figure 4.4 [240]: Sample temperature is an important factor in optimising the conditions for detection of a mutation when using dHPLC. If the temperature is too high, each heteroduplex and homoduplex will be fully denatured and will all elute as single stranded DNA. If the temperatures are too low, each of the heteroduplexes and homoduplexes will be fully annealed and will all elute as double stranded DNA from the column. It is only when the temperature is at a point which will partially denature the region containing the mutation, that heteroduplexes will be detected due to differing ratios of single stranded to double stranded DNA. The wide range of temperatures at which the above mutation in DYS271 was detected (a range of 4°C), meant that a 1°C leeway was allowed for the optimum detection conditions suggested by the melting profiles of each Foxf2 amplicon.

Regions of fragment were therefore considered to carry an acceptable probability of mutation detection when any temperature +/-1°C of the run temperature was predicted to have a helical fraction of 60-99% i.e. all regions which were predicted to be 60-99% helical at either 65, 66 or 67°C would be considered to carry an acceptable probability of mutation detection at 66°C (figure 4.5).

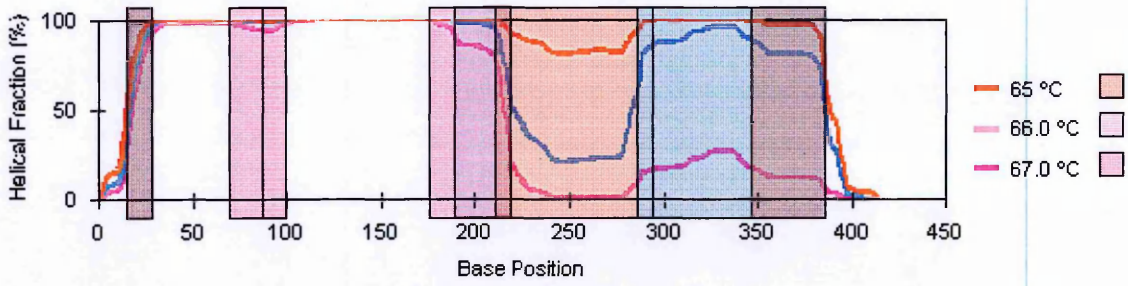


Figure 4.5: Relaxation of dHPLC Temperature Limits
Figure 4.5: Regions of the Foxf2_1_1 amplicon which would be considered to have an acceptable probability of mutation detection at 66°C. Predicted helical fractions of 60-99% for the fragment at 66°C, in addition to fraction percentages of 60-99% at 1°C above and below this temperature, were considered when determining optimum temperatures for screening each fragment.

For each Foxf2 amplicon, these rules were applied in the design of the temperature methods, so that all but the extreme ends of each amplicon had a reasonable chance of mutations being detected along the entire length of each fragment. This is demonstrated in the selected temperatures for Foxf2_1_1 as shown in figure 4.6.

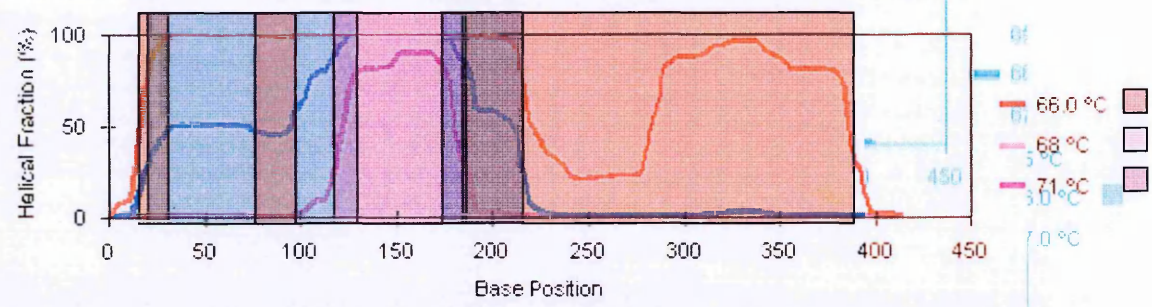


Figure 4.6: Optimal Detection Zones of Foxf2_1_1 at Three dHPLC Temperatures
The regions of Foxf2_1_1 that had an acceptable probability of mutation detection at 66°C, 68°C and 71°C.

4.5 Amplicon Descriptions and dHPLC Runs

The Foxf2 gene sequence was taken from the finished sequence of BAC 467m8 from the RPCI-23 library and submitted to primer3 (http://frodo.wi.mit.edu/) in order to design primers across the gene. Transgenomic recommended that PCR amplicons should be below 450bp for optimum probability of mutation detection, but that

mutations have been detected with fragments as large as 1.5kb [240]. Some of the amplicons designed were over the recommended optimum size (figure 4.7). Following optimisation of PCR products, and selection of those which gave the clearest bands, 5 amplicons were selected with which to screen the GSK archive:

Primer Set	Forward	Reverse	Product Size
Foxf2_1_1	CTCGCCCGATTTGTGGAC	AGCGCGATGTACGAGTAAGG	414
Foxf2_1_1c	AGTGGAGGCACCAAGAAGG	GGAACGAACCCTCCTCAAAC	316
Foxf2_1_2a	TTCCCCTTTTTCCGTGGCGC	TGGCCATATAGGTGGAGCCC	460
Foxf2_1_3a*	TCAAGGCGGTTATGGTGGCC	AGAGGCTCTCAGAGGCTCCG	513
Foxf2_1_4&3a*	ACACCACCTCCACCACCAC	AGAGGCTCTCAGAGGCTCCG	433
Foxf2_2_01	AGCTGCCTTTACACCCTCAG	ACAGTGTGAGTCCGTTGCAG	384

Table 4.1: Primers used to screen the archive.
Primers used for Genotypic Screening of the mutagenised DNA archives.
* Primer set Foxf2_1_4&3a replaced Foxf2_1_3a for the screen of the MRC archive.

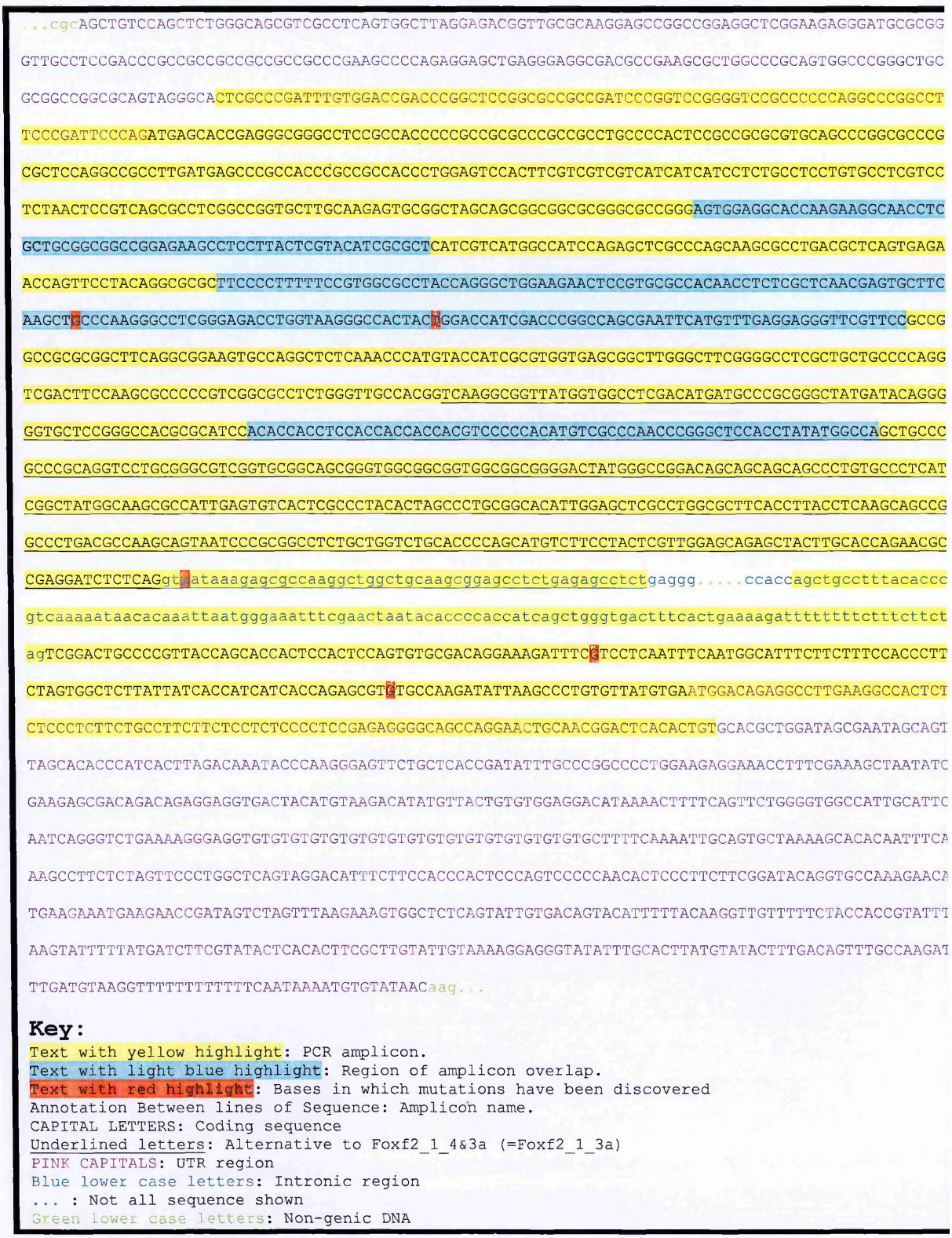


Figure 4.7: Foxf2 transcript
Foxf2 transcript and exon information showing highlighted WAVE and sequencing amplicons.

The first four amplicons: Foxf2_1_1 – Foxf2_1_3a were designed using coding sequence of the first exon of the gene. The 5th amplicon included the entire coding region of the second exon. All primers can be seen in table 4.1

The PCRs were optimised as described in section 2.2.5 and screening of the GSK archive was carried out as described in section 2.2.25. The high GC content of fragment Foxf2_1_1 meant that betaine was required for amplification of this fragment. Although the use of betaine can cause problems with the dHPLC columns, Transgenomic Ltd investigated its use and found that concentrations of under 2M were compatible with the columns. The concentration of betaine used in these PCRs (1.35M) was therefore below the maximum levels recommended by Transgenomic.

After the completion of the GSK archive screen, further optimisation was carried out to increase the strength of signal within each sample. This was necessary due to the high ratio of samples which failed to amplify. Following this additional optimisation, the forward primer which was used for the amplicon Foxf2_1_3a was replaced with an alternative primer to produce a different fragment to cover this region: Foxf2_1_4&3a. This amplicon produced a stronger, cleaner signal than the Foxf2_1_3a product.

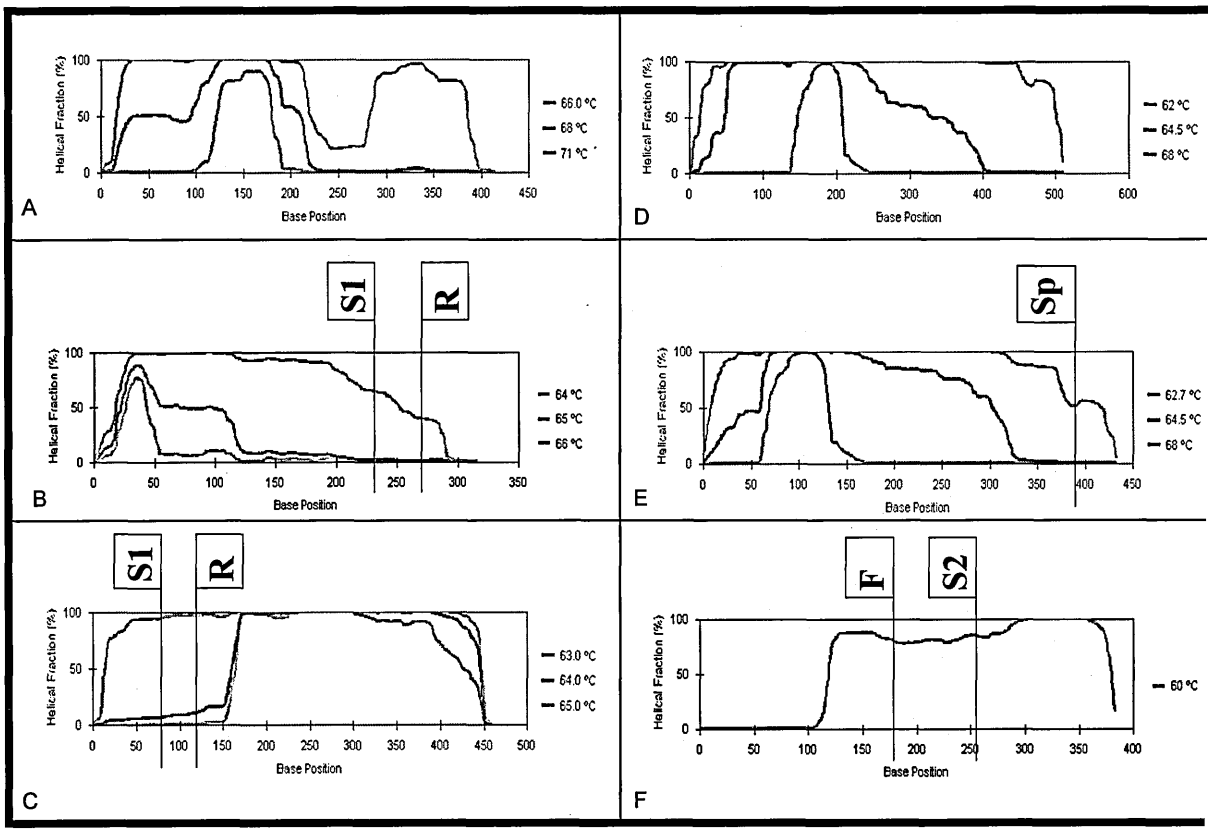


Figure 4.8: Helical fractions for each amplicon
Helical Fractions of the amplicons which were used to screen Foxf2, including the mutations which were discovered and their positions within the fragments.
A. Foxf2_1_1 at 66°C, 68°C and 71°C. **B.** Foxf2_1_1c at 64°C and 65°C including mutations S1 = Silent 1. R = W174R. **C.** Foxf2_1_2a at 63°C, 64°C and 65°C also including mutations S1 and R. **D.** Foxf2_1_3 at 62°C, 64.5°C and 68°C. **E.** Foxf2_1_4&3a at 62.7°C, 64.5°C and 68°C including the mutation Sp = Splice site mutation. **F.** Foxf2_2_1 at 60°C including the mutations S2 = Silent 2 and F = V412F.

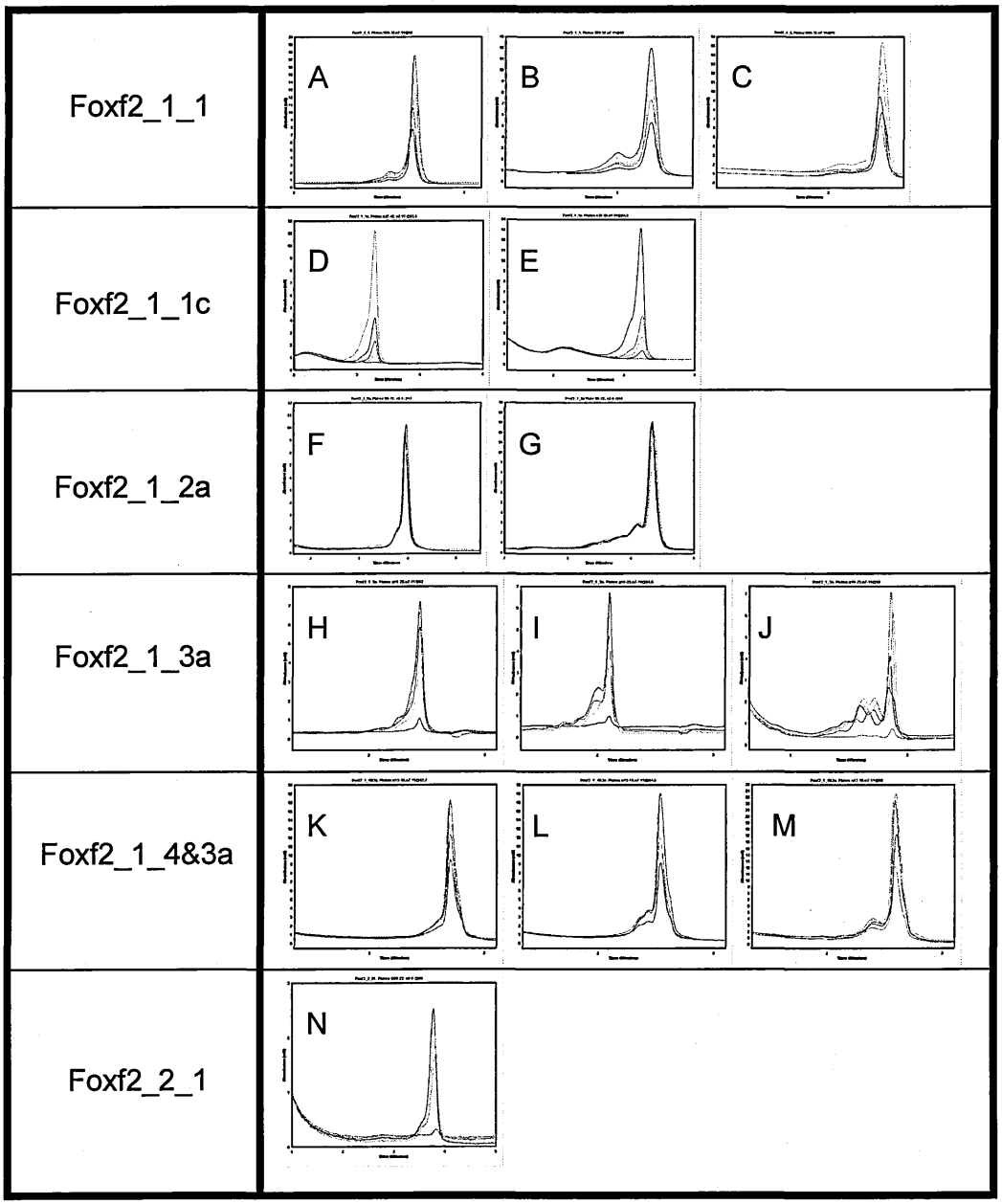


Figure 4.9: dHPLC Electropherograms
The amplicons which were used in the screening of both archives are shown. These are archive samples which did not contain any detectable mutations. Each amplicon is shown at each of the run temperatures used, with the lowest temperatures on the left. A-C: Foxf2_1_1 at 66°C, 68°C and 71°C. D-E: Foxf2_1_1c at 63.9°C and 64.9°C. F-G: Foxf2_1_2a at 63°C and 65°C. H-J: Foxf2_1_3a at 62°C, 64.5°C and 68°C. K-M: Foxf2_1_4&3a at 62.7°C, 64.5°C and 68°C. N: Foxf2_2_1 at 60°C

Mutations

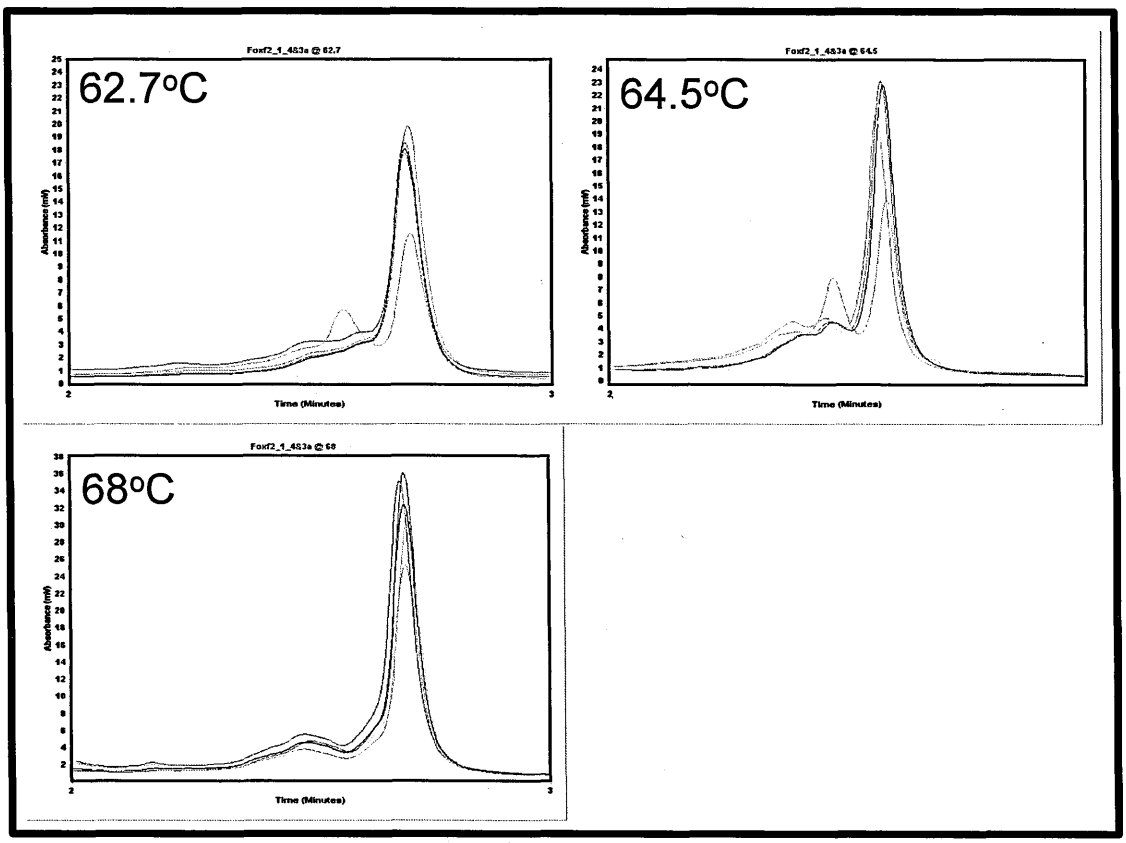


Figure 4.10: Electropherograms for Splice Mutant
Electropherograms from the splice site mutation which was identified in MRC archive sample pool 31H8 with primer set Foxf2_1_4&3a at 62.7oC and 64.5oC. Results from each pool containing a mutant individual is shown superimposed upon 4 other adjacent pooled samples, which act as controls. The mutant pool is shown in orange at each temperature. No difference was detected between the mutant sample and controls at 68oC.

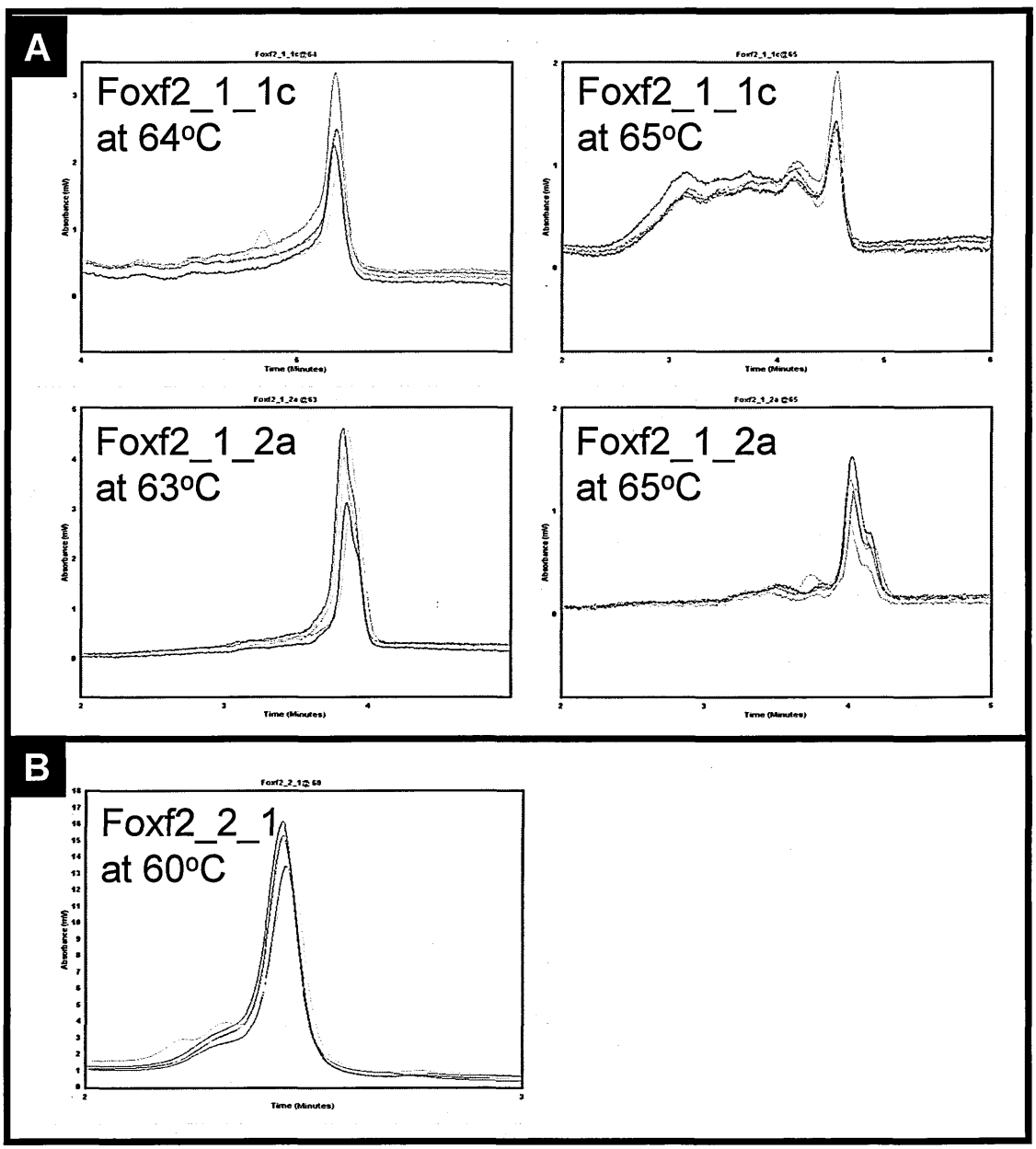


Figure 4.11: Electropherograms for missense mutations
Electropherograms from dHPLC output for the two missense mutations that were identified. Each pool of four samples containing a mutant individual is shown superimposed on 4 other adjacent pooled samples which act as controls. The mutant pool is shown in orange at each temperature. The Missense mutation in mutant A is from the GSK archive in individual 14H3 and was amplified by primer set Foxf2_1_1c at 64°C and 65°C, and primer set Foxf2_1_2a at 63°C and 65°C. The mutation was detected in the Foxf2_1_1c product at 64oC and in the Foxf2_1_2a product at 65oC. Mutation B is from the MRC archive in individual 18c1 and was identified from product amplified by primer set Foxf2_2_1 at 60°C.

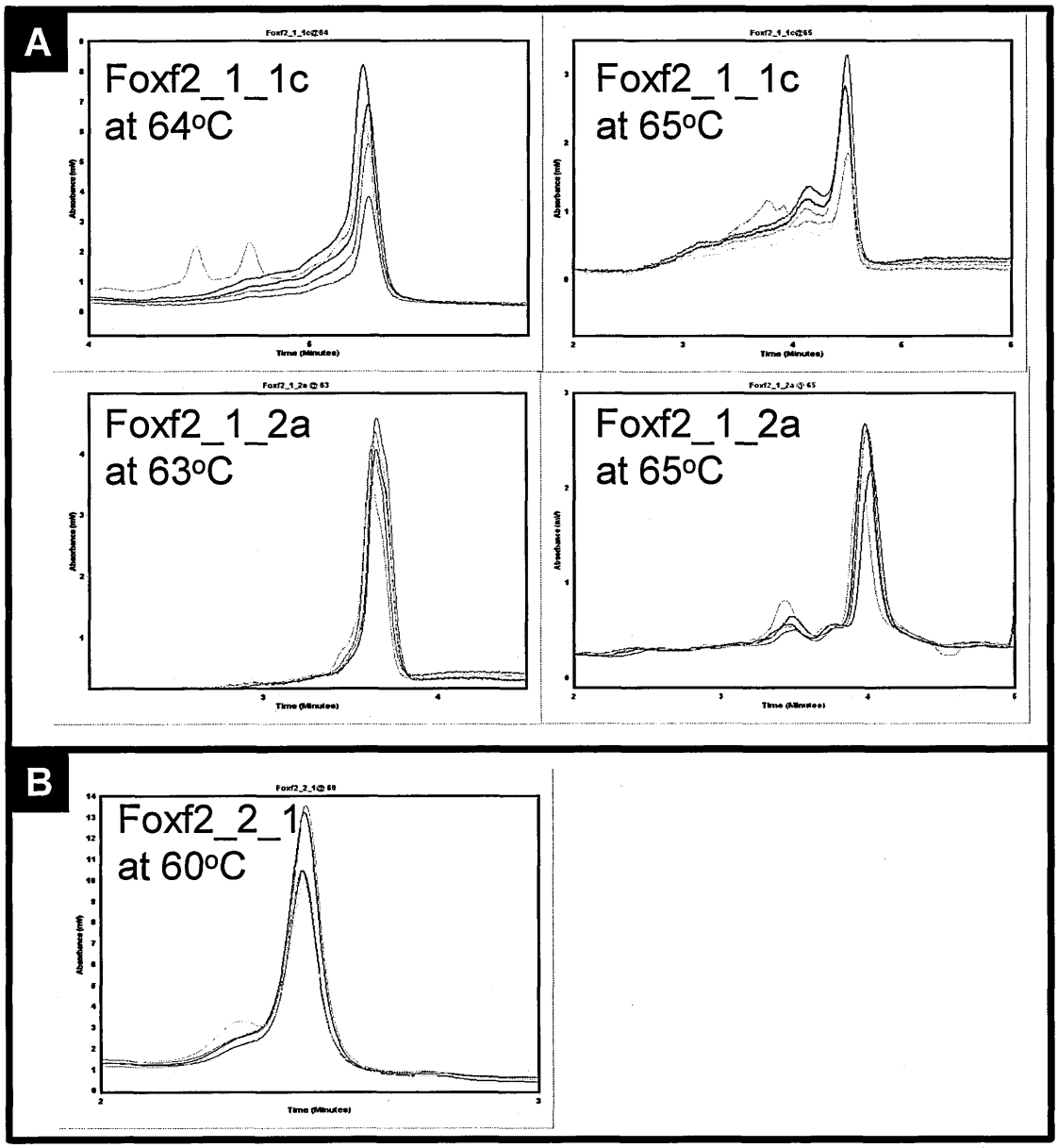


Figure 4.12: Electropherograms for silent mutations
Electropherograms from dHPLC output for the two silent mutations that were identified. Each pool containing a mutant individual is shown superimposed upon 4 other pooled samples which act as controls. A: Silent mutation S1 is from the GSK archive in individual 1B11. This was amplified by primer set Foxf2_1_1c at 64°C and 65°C, and primer set Foxf2_1_2a at 63°C and 65°C. The mutation was detected in the Foxf2_1_1c product at both temperatures and in the Foxf2_1_2a product at 65oC. B: Silent mutation S2 is from the MRC archive at 25H1 and was identified from product amplified with primer set Foxf2_2_1 at 60°C.

4.6 Mutant Peak Analysis

Each of the mutations detected by dHPLC are described by referring to the sample in which they were detected. The complete plate coordinates of each sample have been indicated by the archive source (i.e.: GSK or MRC archives), followed by the plate number then the well reference according to a standard 96 well grid (A-H;1-12). For example, "GSK 1B11" refers to GSK archive plate 1, well co-ordinate B11.

Mutation 1 - GSK 1B11

This mutation was detectable at both 64°C and 65°C with amplicon Foxf2_1_1c (fig4.12 A). The mutation was less clear with the Foxf2_1_2a product than with Foxf2_1_1c at both of the run temperatures for this fragment. A very subtle additional shoulder is present on the peak at 63°C. There is a marked increase in the size of an additional early peak at 65°C.

Each of the DNAs within the pool GSK plate 1-4 B11 were sequenced with primers for Foxf2_1_1c. This showed that individual GSK 1B11 carried a G-T mutation in base 784 of the Foxf2 transcript (fig 4.15A). This base change is a silent mutation, so the leucine for which it codes at amino acid position 161 remains unchanged.

Splice Site Mutation - MRC 31H8

The mutant peak stands out clearly from the other samples at 62.7 °C and 64.5 °C but cannot be differentiated from the other samples at 68°C (fig 4.10).

Each of the DNAs within the pool MRC plate 29-32 H8 were sequenced with primers for Foxf2_1_4&3a. The sequence showed that individual MRC 31H8 carried an A-G mutation in base 3 of the only intron within Foxf2 (fig 4.15C).

V-F mutation - MRC 18C1

Two additional peaks prior to the main peak were present in individual MRC 18C1which could not be seen on adjacent samples, suggesting the presence of a mutation (fig 4.11 B). Each of the DNAs within the pool MRC plate 17-20 C1 were sequenced with Foxf2_2_1. The sequence showed that individual MRC 18C1 carried a G-T mutation in base 1535 of the Foxf2 transcript(fig 4.15D) which would result in an amino acid change from a valine to a phenylalanine at position 412 of the protein.

W174R mutant peak - GSK: 14H3

A sharp, distinctive additional peak prior to the main elution peak could be seen at 64°C but not at 65°C with amplicon Foxf2_1_1c (fig 4.11 A). This sample and those around it have produced low product yields, and the peaks seem particularly small at 65°C. There is a high background signal in the region prior to the main peak at 65°C.

Despite low absorbance levels of under 2mV in the W174R sample and adjacent samples with Foxf2_1_2a at 65°C, the mutation is clearly visible. There is no difference between the mutant and surrounding peaks when Foxf2_1_2a was used to screen the W174R mutant at 63°C (fig 4.11 A).

Each of the DNAs within the pool GSK plate 13-16 H3 were sequenced with both Foxf2_1_1c and Foxf2_1_2a primer sets. The cleanest sequence was from Foxf2_1_2a – this showed that individual GSK 14H3 carried a T-A mutation in base 821 of the Foxf2 transcript (fig 4.15E) which would result in an amino acid change from a tryptophan to an arginine at amino acid position 174 of the protein. Additional proof of the mutation was gained from Sfc1 digestion of the PCR product (figure 4.13). No Sfc1 sites exist within the wildtype amplicon, but this mutation would create one (figure 4.14).

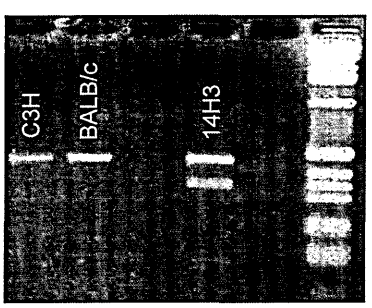


Figure 4.13: SfcI mutation confirmation of W174R mutation
The double band on the right is Sfc I digested Foxf2_1_2a PCR fragment from a heterozygous W174R mutant individual. The Two single bands are C3H & BALB/c respectively – these remained undigested in the presence of Sfc I.

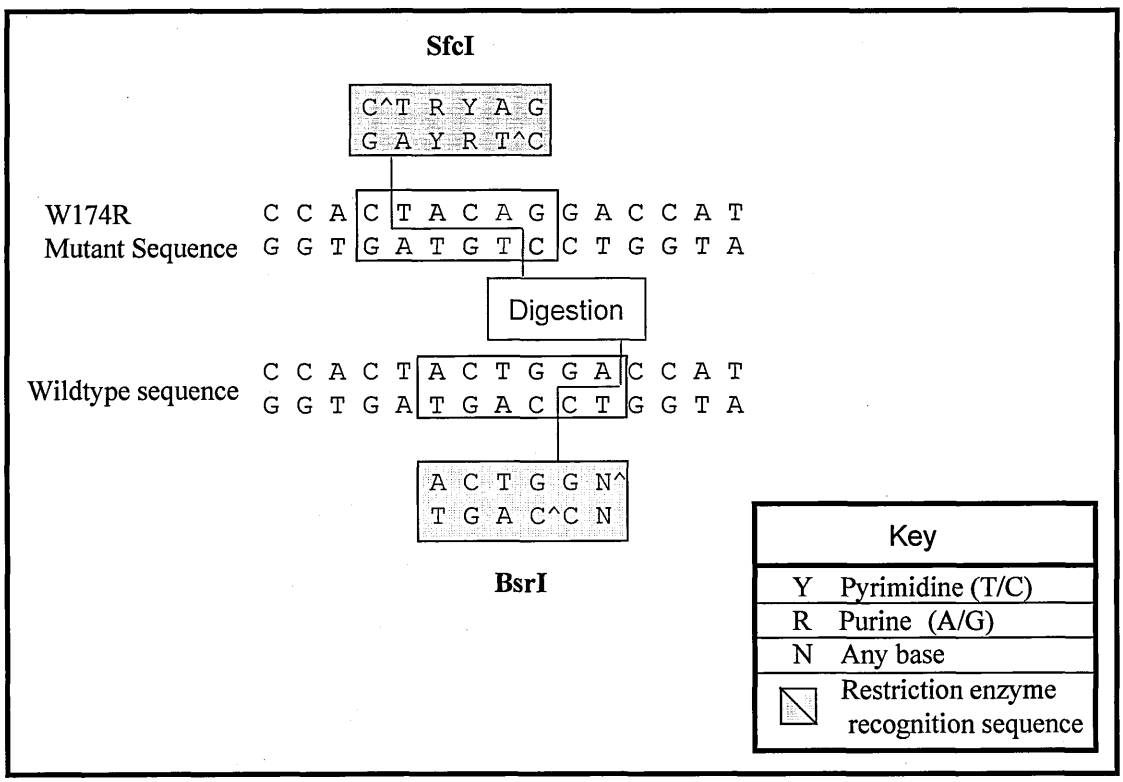


Figure 4.14: Restriction sites in Foxf2 W174R mutant
The A to T point mutation in the W174R mutant results in the removal of a BsrI recognition sequence and its replacement with an SfcI site. The mutant base pair is shown in red.

Mutation Silent 2 - MRC: 25H1 (MRC plate 25, co-ordinate H1)
This mutation was the second to be detected within amplicon Foxf2_2_1, and as such, was only run at one temperature: 60°C (fig 4.12, image B). A significant additional peak can be seen prior to and within the shoulder region of the main elution peak.

Each of the DNAs within the pool MRC plate 25-28 H1 were sequenced with Foxf2_2_1.This showed that individual GSK 25H1 carried a G-C mutation in base 1575 of the Foxf2 transcript (fig 4.15B). This base change is a silent mutation, so the valine for which it codes at amino acid position 437 remains unchanged.

.

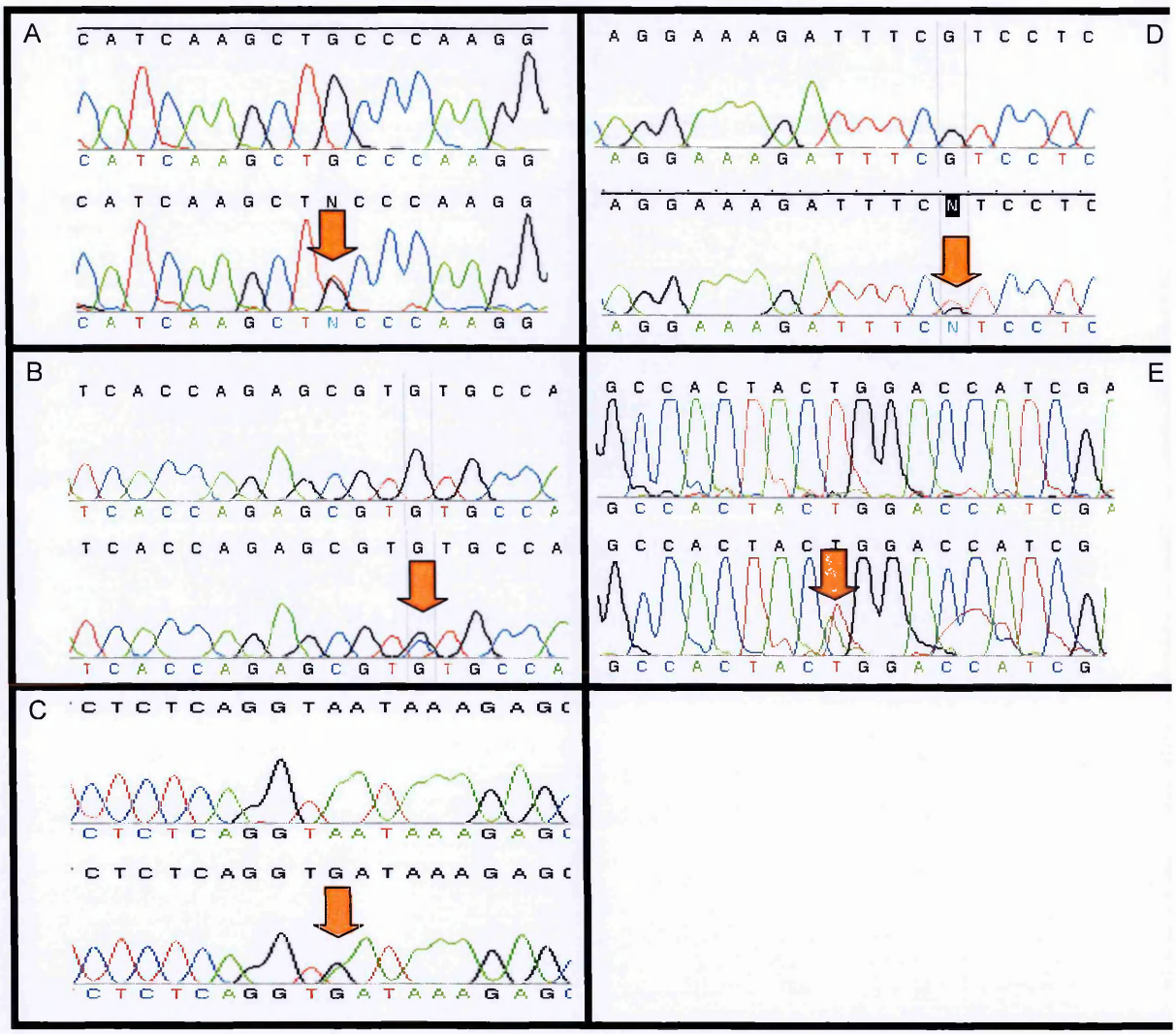


Figure 4.15: Mutation sequencing
Sequence chromatograms of each of the 5 mutations discovered in Foxf2. In each example, wildtype control sequence (C3H or BALBc – both of which carry the same sequence throughout Foxf2) is shown at the top and the mutant sequence at the bottom. Red arrows indicate the precise position of each mutation. **A**. Silent Mutation 1 within the amplicon Foxf2_1_1c in mouse (GSK ID) 1B11. **B**. Silent Mutation 2 within the amplicon Foxf2_2_1 in mouse MRC 25H1**. C**. Splice Site Mutation within amplicon Foxf2_1_4&3a in mouse (MRC archive ID) 31H8**. D**. V412F Mutation within the amplicon Foxf2_2_1 in mouse (MRC archive ID) 18C1. **E**. W174R Mutation within the amplicon Foxf2_1_2a in mouse (GSK ID) 14H3.

```
ATG AGC ACC GAG GGC GGG CCT CCG CCA CCC CCG CCG CGC CCG CCG CCT GCC CCA CTC CGC CGC GCG TGC AGC
Met Ser Thr Glu Gly Gly Pro Pro Pro Pro Pro Arg Pro Pro Pro Ala Pro Leu Arg Arg Ala Cys Ser

CCG GCG CCC GGC GCG CTC CAG GCC GCC TTG ATG AGC CCG CCA CCC GCC GCC ACC CTG GAG TCC ACT TCG TCG
Pro Ala Pro Gly Ala Leu Gln Ala Ala Leu Met Ser Pro Pro Pro Ala Ala Thr Leu Glu Ser Thr Ser Ser

TCG TCA TCA TCA TCC TCT GCC TCC TGT GCC TCG TCC TCT TCT AAC TCC GTC AGC GCC TCG GCC GGT GCT TGC
Ser Ser Ser Ser Ser Ser Ala Ser Cys Ala Ser Ser Ser Ser Asn Ser Val Ser Ala Ser Ala Gly Ala Cys

AAG AGT GCG GCT AGC AGC GGC GGC GCG GGC GCC GGG AGT GGA GGC ACC AAG AAG GCA ACC TCG GGG CTG CGG
Lys Ser Ala Ala Ser Ser Gly Gly Ala Gly Ala Gly Ser Gly Gly Thr Lys Lys Ala Thr Ser Gly Leu Arg

CGG CCG GAG AAG CCT CCT TAC TCG TAC ATC GCG CTC ATC GTC ATG GCC ATC CAG AGC TCG CCC AGC AAG CGC
Arg Pro Glu Lys Pro Pro Tyr Ser Tyr Ile Ala Leu Ile Val Met Ala Ile Gln Ser Ser Pro Ser Lys Arg

CTG ACG CTC AGT GAG ATC TAC CAG TTC CTA CAG GCG CGC TTC CCC TTT TTC CGT GGC GCC TAC CAG GGC TGG
Leu Thr Leu Ser Glu Ile Tyr Gln Phe Leu Gln Ala Arg Phe Pro Phe Phe Arg Gly Ala Tyr Gln Gly Trp

                                                                    (T)
AAG AAC TCC GTG CGC CAC AAC CTC TCG CTC AAC GAG TGC TTC ATC AAG CTG CCC AAG GGC CTC GGG AGA CCT
Lys Asn Ser Val Arg His Asn Leu Ser Leu Asn Glu Cys Phe Ile Lys Leu Pro Lys Gly Leu Gly Arg Pro
                                                                  (Leu)
                    (A)
GGT AAG GGC CAC TAC TGG ACC ATC GAC CCG GCC AGC GAA TTC ATG TTT GAG GAG GGT TCG TTC CGC CGC CGG
Gly Lys Gly His Tyr Trp Thr Ile Asp Pro Ala Ser Glu Phe Met Phe Glu Glu Gly Ser Phe Arg Arg Arg
                   (Arg)

CCG CGC GGC TTC AGG CGG AAG TGC CAG GCT CTC AAA CCC ATG TAC CAT CGC GTG GTG AGC GGC TTG GGC TTC
Pro Arg Gly Phe Arg Arg Lys Cys Gln Ala Leu Lys Pro Met Tyr His Arg Val Val Ser Gly Leu Gly Phe

GGG GCC TCG CTG CTG CCC CAG GGC TTC GAC TTC CAA GCG CCC CCG TCG GCG CCT CTG GGT TGC CAC GGT CAA
Gly Ala Ser Leu Leu Pro Gln Gly Phe Asp Phe Gln Ala Pro Pro Ser Ala Pro Leu Gly Cys His Gly Gln

GGC GGT TAT GGT GGC CTC GAC ATG ATG CCC GCG GGC TAT GAT ACA GGG GCG GGT GCT CCG GGC CAC GCG CAT
Gly Gly Tyr Gly Gly Leu Asp Met Met Pro Ala Gly Tyr Asp Thr Gly Ala Gly Ala Pro Gly His Ala His

CCA CAC CAC CTC CAC CAC CAC CAC GTC CCC CAC ATG TCG CCC AAC CCG GGC TCC ACC TAT ATG GCC AGC TGC
Pro His His Leu His His His His Val Pro His Met Ser Pro Asn Pro Gly Ser Thr Tyr Met Ala Ser Cys

CCG GTG CCC GCA GGT CCT GCG GGC GTC GGT GCG GCA GCG GGT GGC GGC GGT GGC ATG GCA AGC GCC GGC GGG
Pro Val Pro Ala Gly Pro Ala Gly Val Gly Ala Ala Ala Gly Gly Gly Gly Gly Gly Gly Asp Tyr Gly Pro

GAC TAT GGG CCG GAC AGC AGC AGC AGC CCT GTG CCC TCA TCC CCG GCT ATT GAG TGT CAC TCG CCC TAC ACT
Asp Ser Ser Ser Ser Pro Val Pro Ser Ser Pro Ala Met Ala Ser Ala Ile Glu Cys His Ser Pro Tyr Thr

AGC CCT GCG GCA CAT TGG AGC TCG CCT GGC GCT TCA CCT TAC CTC AAG CAG CCG CCT GCC CTG ACG CCA AGC
Ser Pro Ala Ala His Trp Ser Ser Pro Gly Ala Ser Pro Tyr Leu Lys Gln Pro Pro Ala Leu Thr Pro Ser

AGT AAT CCC GCG GCC TCT GCT GGT CTG CAC CCC AGC ATG TCT TCC TAC TCG TTG GAG CAG AGC TAC TTG CAC
Ser Asn Pro Ala Ala Ser Ala Gly Leu His Pro Ser Met Ser Ser Tyr Ser Leu Glu Gln Ser Tyr Leu His

                                    (g)
CAG AAC GCC CGC GAG GAT CTC TCA Ggtaa...cagTC GGA CTG CCC CGT TAC CAG CAC CAC TCC ACT CCA GTG TGC
Gln Asn Ala Arg Glu Asp Leu Ser Val(Intron) Gly Leu Pro Arg Tyr Gln His His Ser Thr Pro Val Cys

                   (T)
GAC AGG AAA GAT TTC GTC CTC AAT TTC AAT GGC ATT TCT TCT TTC CAC CCT TCC GCT AGT GGC TCT TAT TAT
Asp Arg Lys Asp Phe Val Leu Asn Phe Asn Gly Ile Ser Ser Phe His Pro Ser Ala Ser Gly Ser Tyr Tyr
                   (Phe)
                         (C)
CAC CAT CAT CAC CAG AGC GTG TGC CAA GAT ATT AAG CCC TGT GTT ATG TGA
His His His His Gln Ser Val Cys Gln Asp Ile Lys Pro Cys Val Met STOP
                      (Val)
```

Figure 4.16: Foxf2 Mutations in relation to the transcript
Each mutation within the transcript and abbreviated intron (red text). The mutation positions have been highlighted within the wildtype transcript sequence and mutant bases and amino acids are annotated in brackets.

4.7 Mutation Discovery Discussion

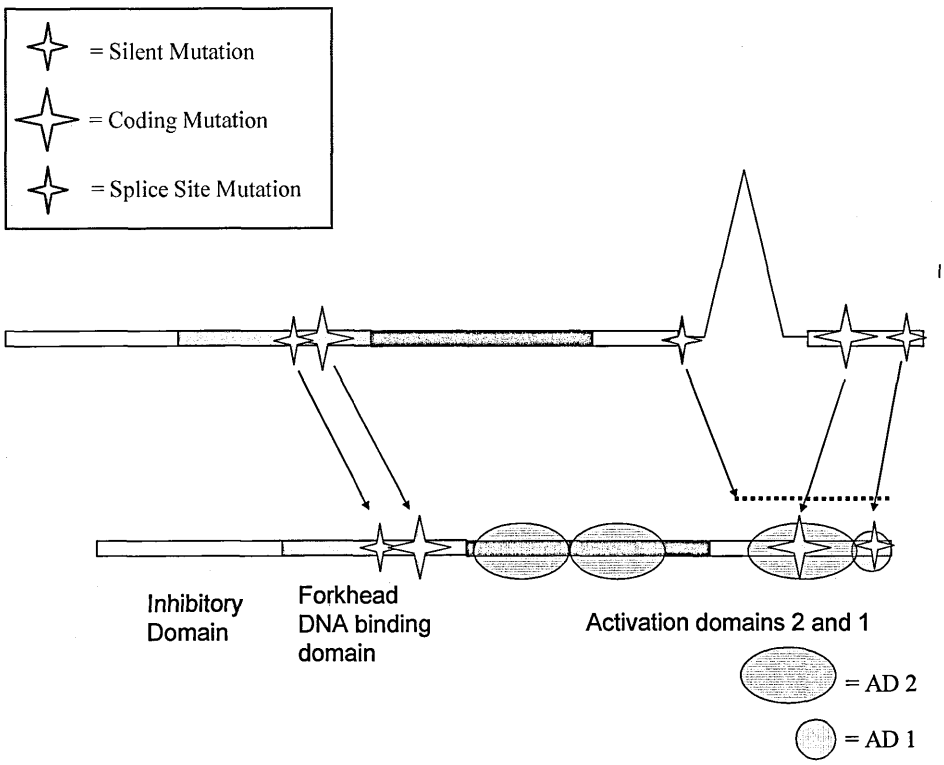


Figure 4.17: Foxf2 mutations in relation to the functional domains
Illustration of the four known functional domains of Foxf2 including the approximate locations of mutations which were discovered with the gene driven screen. The domain structure is shown as described in the mouse study by Miura [241]. These activation domains are shown overlayed by the synergistic sub-domains that were described for the human gene [242]. Sequence identity between the human and mouse transcript is ~89% and this suggests a high degree of functional conservation, which is why it was considered reasonable to combine mouse and human domain studies. The non-synonymous mutations have been represented as stars in both the DNA and protein images, the splice site mutation has been indicated with a dotted line above the region of protein derived from exon 2, representing an unknown splicing efficiency.

G-T mutation (Silent 1) at position 231 GSK archive 1B11

This mutation was obvious in Foxf2_1_1c at both 64°C and 65°C (fig 4.12A), supporting the practice of allowing 1°C leeway in the design of temperature methods for each fragment. The mutation was less obvious in the Foxf2_1_2a product in which the mutation was also located (in the overlapping region of the two products). The very subtle additional shoulder which could be seen at 63°C was unlikely to have been recognised without prior knowledge of the increased probability of a

mutation in this sample. The increased resolution of the mutant peak at 65°C is a surprise given the low level of the peak at this temperature, especially when the helical fraction prediction indicates that mutation detection is more likely at 63°C (fig 4.8C). (See also section **Foxf2_1_2a with W174R and Silent 1)**

Splice Site Mutation. MRC archive **31H8**

That this mutation is detectable at 62.7°C is not surprising given that the helical fraction prediction for this position at 62.7°C lies at around 50% (fig 4.8E). It is of interest, however, that the mutation was even more clearly differentiated at 64.5°C when the predicted helical fraction for the position at this temperature was 0%. This gives further credence to the practice of allowing a 1°C leeway when designing temperature methods to achieve helical fractions of between 60 and 99%.

This mutation occurs at the 3rd base after the end of the 1st exon (fig 4.18). It could be regarded as a conservative change as approximately 35% of splice sites contain a G at this position [243]. Each of the other bases within the splice region are much more highly conserved within splice sites than this one, indicating that this mutation at this locus is less likely to have a severe affect on splicing efficiency than would other mutant forms within this seven base region.

There has been one recorded case in which an identical base change in splice site donor sequence (A-G at intron position +3) has affected splicing however. This occurred at the 5' end of intron 7 of ornithine transcarbamylase (OTC) and caused

the excision of the preceding exon (exon 7), resulting in OTC deficiency and subsequent ammonia intoxication [244]. Carstens proposed two potentially compounding reasons for the failure of correct splicing. The proposal hinges on a model that was previously suggested by Robberson for the recognition of exons by the spliceosome: the spliceosome pre-defines the exon before the commencement of splicing by first recognising the acceptor site at the 3' end of one intron, and then the downstream donor site at the 5' end of the next intron [245]. Carstens hypothesised that the efficiency of this exon recognition may be reduced by the base change. He also suggested that the short length of the Intron (80 nucleotides) may have further reduced splicing efficiency resulting in the removal of exon 7.

The environment of the OTC splice site that was of interest to Carstens is very different from that seen in the mouse Foxf2 mutation. In Foxf2 for example, the recognition of the first donor sequence in a gene is likely to be mediated by different factors than those bordering internal exons because there is no upstream 3' acceptor. Another difference is that the Foxf2 intron length (3.2kb in mouse and 3.5kb in human) easily exceeds the minimum length for correct splicing activity. It is therefore difficult to predict the potential consequences of this splice site mutation in Foxf2. The longer intron suggests that the recognition sequence may not be as critical because intron length is not limiting splicing efficiency. However the OTC mutation provides evidence that such a mutation could affect splicing efficiency, making this mutation worth further investigation, not least because it may provide an

insight into the hypotheses which were used to explain the OTC mutant splice pattern.

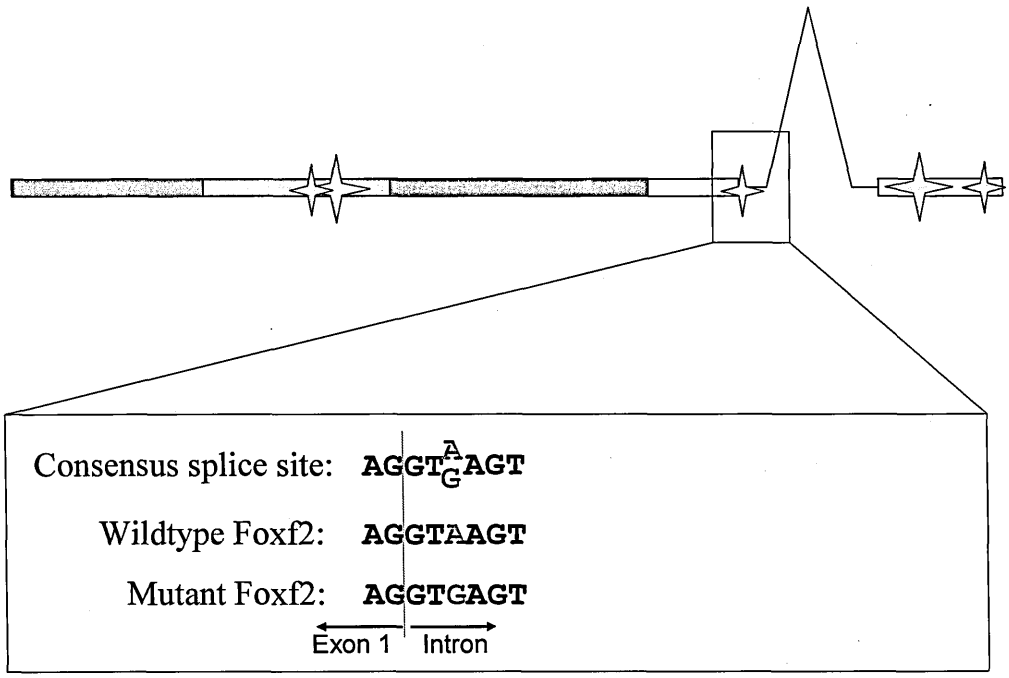


Figure 4.18: Splice site mutation in relation to the exon/intron boundary
The intron/exon boundary is indicated by a vertical orange line within the sequence. The mutation occurs on the third base into the intron and is a G-A mutation. Approximately 60% of splice sites have a G at this position and 35% have an A, whereas none of the other positions contain a secondary base at such a high frequency ([243]). This suggests that the mutation may not have a significant effect on splicing efficiency.

V-F mutation. MRC18c1

The position of the valine to phenylalanine mutation is predicted to fall within the optimum helical fraction parameters for mutation detection at the run temperature of 60°C. The peaks were clear and the mutation was easily detected as would be expected from the helical fraction prediction data.

The mutated residue is located within the most C terminal sub-domain of AD2 (fig 4.17 and 4.19). This is the same domain, although different sub-domain, in which a glycine insertion within a poly-glycine site in a human family of glaucoma sufferers was identified (C. Gregory Evans, unpublished data). It has been suggested that the

subdomains act synergistically with each other to allow full transcriptional activity [242]. Both valine and phenylalanine are hydrophobic although phenylalanine is hydrophobic to a much greater degree. Valine has an aliphatic side chain whereas phenylalanine is aromatic, so the sizes of the two amino acids are different and this could affect the structure of the protein.

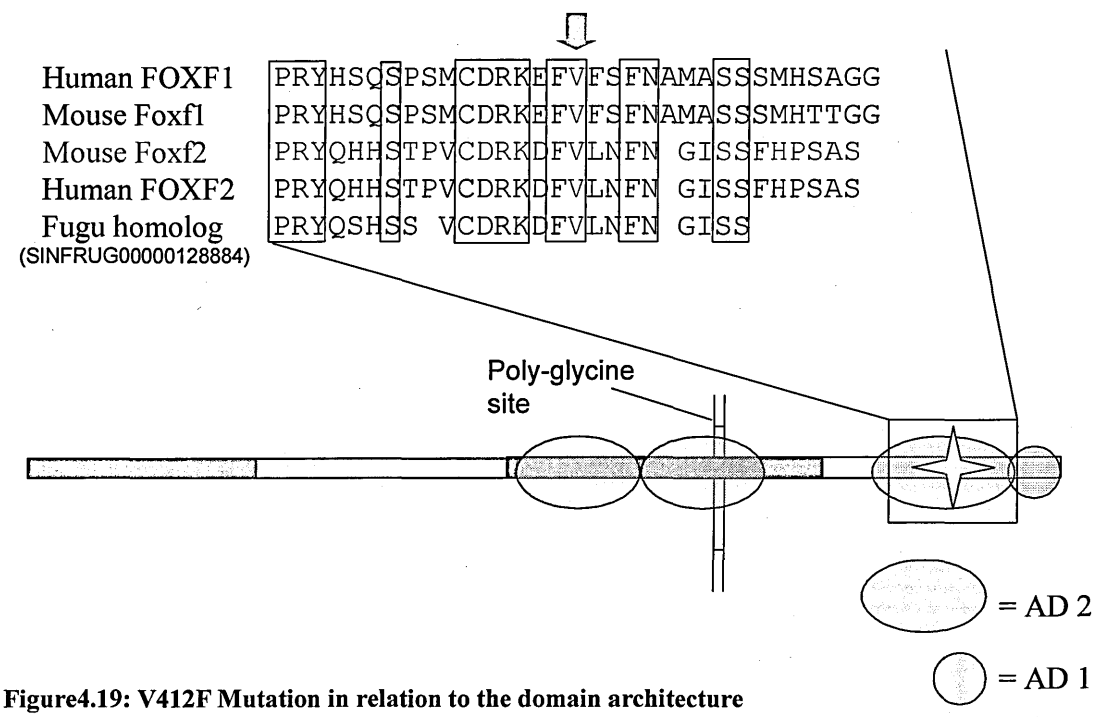


Figure4.19: V412F Mutation in relation to the domain architecture
The valine at residue 412 of Mouse Foxf2 is conserved throughout the Foxf clade in vertebrates. It is located within a subdomain of activation domain 2. This is the same domain, although different subdomain, from the location in which a glycine insertion into a polyglycine repeat was identified in a family of glaucoma sufferers.

Silent 2 Mutation (MRC 25H1)

This peak was clearly different from other samples as would be expected from a mutation at this point within the fragment, according to the helical fraction prediction of between 60-99% at the 60°C run temperature.

W174R Mutation (GSK 14H3)

The poor quality and low yield (less than 2mV) of this product seen with amplicon Foxf2_1_1c at 65°C is likely to be obscuring evidence of potential heteroduplexes. It is surprising that the mutation was not detectable in the Foxf2_1_2a product at 63°C when it was distinguishable at 65°C. The helical fraction at the region in which the mutation occurs is predicted to be just under 100% at 63°C (fig 4.8c). 63°C looks to be the most suitable temperature possible for mutation detection in this region of the fragment. In contrast to this, at 65°C the predicted helical fraction is 0%, which would suggest that all of the DNA in the sample at this region would be single stranded DNA and therefore not likely to produce heteroduplexes.

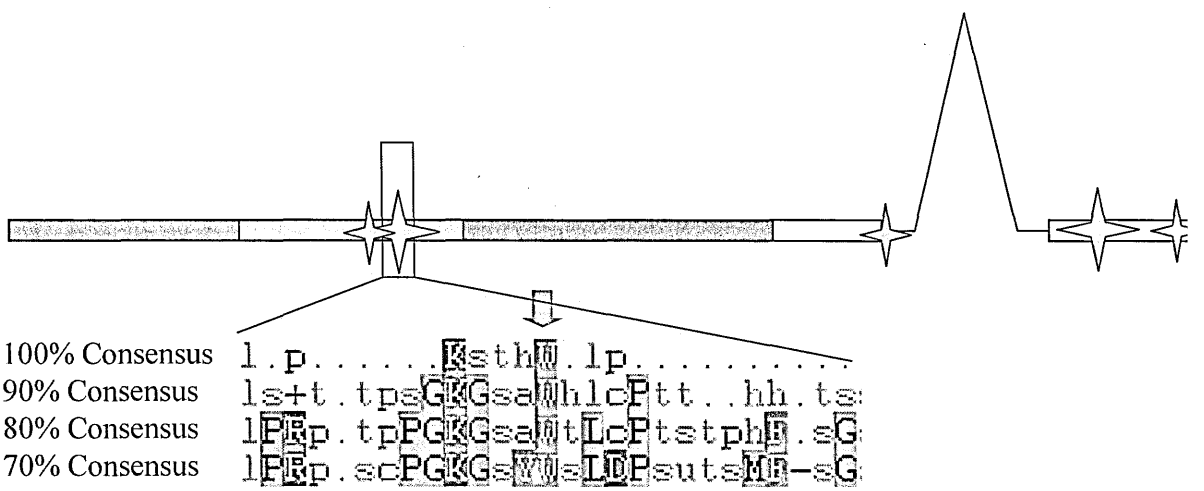


Figure 4.20: W174R mutation in relation to the forkhead domain
The W174R mutation occurs within the DNA binding (forkhead) domain of Foxf2. W174 is a very highly conserved residue which can be seen in every known member of the forkhead family. Tryptophan (W) has different physical properties from arginine (R).

Foxf2_1_2a with W174R and Silent 1

An unexpected result in both the W174R mutation and the Silent 1 mutation, was that the differentiation of mutant samples was clearer at 65°C than at 63°C. An important premise of the technique is that mutations reduce the ratio of double stranded DNA when temperatures are at partial denaturing levels. With predicted helical fractions at 0% for both mutation sites at 65°C, I would have expected that if one of the temperatures was to produce results with improved mutation resolution, it was likely to be at 63°C. This tryptophan to arginine mutation was the first Foxf2 mutation found within the GSK archive. It occurs within the DNA binding (forkhead) domain in a residue which is conserved in all family members (fig4.20).

These two amino acids have very different physical properties from each other. Tryptophan has a neutral pH, is hydrophobic and has a neutral charge whereas arginine is basic and hydrophilic with a positive charge. Because the mutation occurred within the DNA binding domain at an extremely highly conserved position, we believed that this significant amino acid change could cause disruption to the activity of the Foxf2 protein.

The mutated tryptophan locus is within a β-sheet of the protein (fig4.21). This is predicted to be completely disrupted by the mutation according to results from the secondary structure prediction programme DSC, and shortened with additional downstream effects by another programme (Simpa96) as shown in figure 4.22.

NORSp (Non regular secondary structure prediction) written by Jinfeng Liu & Burkhard Rost [246] predicted no affect on the β-sheet containing the mutant residue, but did predict the alteration of the exposure/burial status to a wide range of residues across the protein, and the complete disruption of two strand structures at the carboxyl end of the protein (figure 4.21 C).

Consolidation of these secondary structure predictions is difficult due to the apparent inconsistency between the output of each software – there is not conclusive evidence that the b-sheet containing the mutant residue is altered or disrupted. However, each of the programmes predict the presence of the β-sheet structure surrounding the tryptophan residue in the wildtype and all of them suggest an alteration to the basic protein structure in the mutant, albeit to varying degrees of severity.

PIX was a programme hosted by the Human Genome Mapping Project Resource Centre (HGMP-RC) at www.hgmp.mrc.ac.uk/, which is no longer operational. DSC itself was also hosted by the HGMP and is also therefore individually unavailable, however, Simpa96 (*http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_simpa96.html)* and Predator (*http://www.embl-heidelberg.de/cgi/predator_serv.pl)* are both independently available.

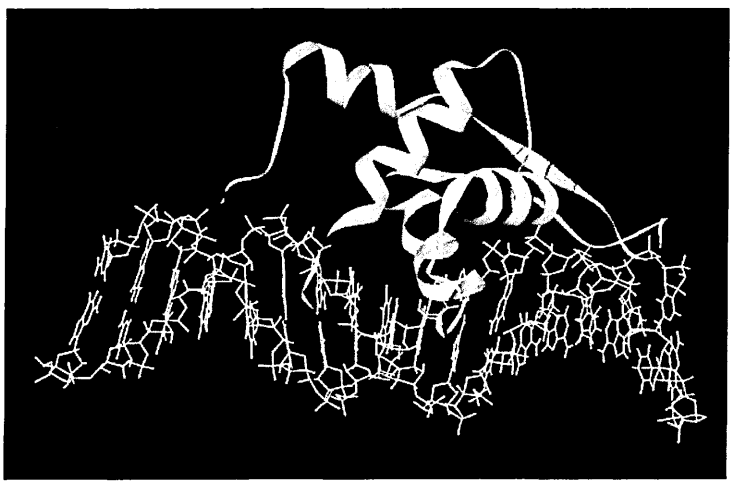


Figure 4.21: W174R mutation & 3D structure of a related forkhead domain
The DNA binding domain of Rat *Foxd3* as it binds to DNA[247]: There is a 62.9% sequence identity between the binding domains of *Foxf2* and *Foxd3* (http://expasy.org/tools/sim-prot.html). The physical structure of the domain is thought to be well conserved throughout the gene family. The Tryptophan residue which is mutated in the W174R mutation in conserved in *Foxd3* and is highlighted in yellow within one of the beta sheets (shown here in pink). The golden coil structures are alpha helices. This diagram was produced by Protein Explorer using accession number 2HDC (http://molvis.sdsc.edu/protexpl/frntdoor.htm).

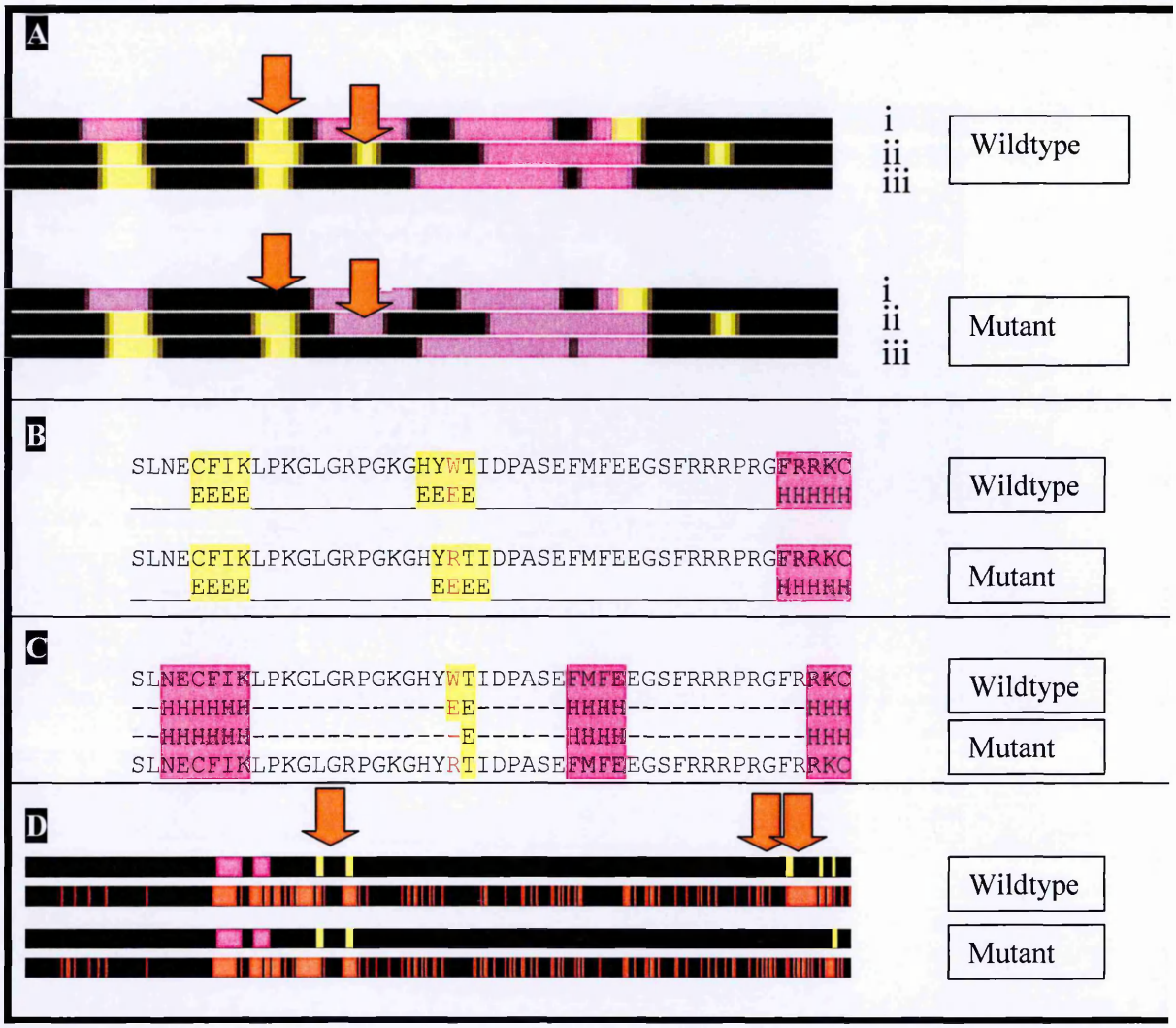


Figure 4.22: Effect of W174R on secondary structure predictions
Secondary structure was predicted by a range of different software. (A) Data from three separate programmes incorporated into the PIX analysis suite; each of the three programmes is represented as a separate horizontal strip for a region of the protein surrounding both the wildtype and mutant alleles. β-sheets are shown as yellow bands and alpha helices as pink bands. Disruption of a beta sheet structure at the mutant locus (arrow 1), is predicted in the first programme (i: DSC). In a second programme (ii: Simpa96), the beta sheet containing the mutation is shortened and, downstream, a different beta sheet region is predicted to be disrupted by the formation of a novel alpha helix in the mutant form of the protein (arrow 2). The third programme (iii: PREDATOR) did not show any affect at the mutant residue although a downstream a-helix was extended. However, a more recent version of this programme when run independently (B) showed that the b-sheet structure containing the mutant residue shifted one position towards the carboxyl terminus of the protein in the mutant. The programme nnpredict (C) predicted a small strand structure at the tryptophan residue at position 174 and the adjacent threonine at 175, that is reduced to only the threonine residue in the mutant, whereas the rest of the protein remains unaffected. Analysis of the whole protein by NORSp (D) was converted into the same graphical format as the PIX programmes, the secondary structure (pink and yellow bands = helices and strand structures respectively) and exposure/burial status of each residue is shown (red = buried, black = exposed). The site of the W174R mutation is indicated by arrow 3 – no alteration of the secondary structure at the mutant loci or the region immediately surrounding it is

predicted. However, the exposure status of other residues from across the protein have been altered, with the most significant effect occurring near to the carboxyl terminus, where two strand structures have been disrupted in the mutant (arrows 4 &5). Both nnpredict and NORSp were accessed via the PredictProtein–Meta portal (http://www.predictprotein.org/meta.php).

Mutation Detection Probabilities

During this screen, 5 mutations were discovered from a screen of 10.98Mb of total DNA screened, this equates to 1 mutation per 2.2Mb. Three of these mutations were potentially functional which corresponds to 1 potentially functional mutation per 3.66Mb. Two were coding mutations, giving a rate of 1 coding mutation per 5.49Mb.

Mutation rates from the combined data of 15 gene driven screens, including the GSK data for Foxf2, found a mutation rate of 1 in 1.01Mb, of which 1 in 1.82Mb was coding [202]. The rates of detection of Foxf2 was therefore lower than those of other screens of the same archive. However, some natural variation is likely in any stochastic process and this variation could be explained by chance.

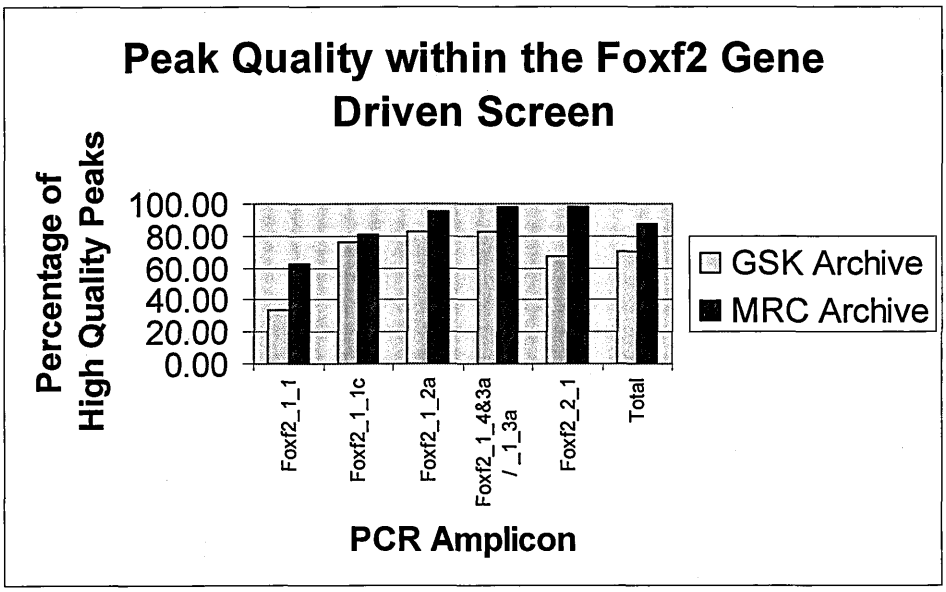


Figure 4.23: Relative PCR Peak quality during archive screening
Relative Peak Quality for each amplicon within both archives used for screening Foxf2. A Significant increase in Peak quality can be seen, following further optimisation after screening the GSK archive.

There are a range of factors which may have led to this outcome. The sub-optimal PCR conditions which were being used for the GSK archive (see figure 4.23) may have caused mutations to remain undetected. The low percentage of good quality peaks for amplicon Foxf2_1_1 during the screening of the GSK archive in particular, may have reduced the possibility of mutation detection in this region – and no mutations were detected within this fragment. However, when the statistics for each archive were analysed individually, the mutation detection rate for the MRC archive (1 mutation per 2.4Mb) was lower than for the GSK archive (1 per 1.9Mb).

It is interesting to compare these figures with those of the larger study. However such a comparison cannot be used to reliably assess the presence of local mutability

variations or sub-optimal PCR conditions, because the number of individuals in each archive is too small for a statistically robust comparison.

The Mutations

Three potentially functional mutations were discovered with the gene driven screen. Judging by the level of conservation between species and the differing physical properties of mutant and wildtype amino acids, the Foxf2W174R mutation looked the most likely to disrupt the function of the gene product. This line was therefore rederived for further analysis.

The V410F mutation and splice site mutation are also of great potential interest due partly because they may provide phenotypes of reduced severity to other Foxf2 mutations, although there is also the distinct possibility that neither would impinge on gene function enough to produce a testable mouse phenotype. Valine and phenylalanine are both hydrophobic and neutral, but with differing levels of hydrophobicity and different side chains. Another factor that suggests the severity of this mutation might not be as great as the W174R mutation is its position within the domain architecture. Despite the relative similarity of the physical properties of valine and phenylalanine, a comparison of various amino acid substitution matrices [248] appeared to show that a replacement of phenylalanine for valine or vice-versa is not a common event – suggesting that this is a non-conservative mutation and its effect may be significant.

W174R is located on a predicted β-sheet of the forkhead domain, and therefore an alteration at this locus may disrupt a critical element to the structure of the region that is responsible for DNA binding. V412F is not positioned within any predicted regular structures but does fall within an activation domain that has been shown to act synergistically with two other domains [242]. The nature of the relationship between these domains has not been investigated further since Hellqvist carried out this preliminary investigation of truncated FOXF2 peptides. The V412F mutation provides a potential opportunity to do this, and it would be particularly interesting to view its activity in comparison with the polyglycine mutation that was identified in glaucoma patients –in one of the other synergistic activation domains [155].

One powerful experiment for the investigation of mutations in the activation domains of transcription factors is the transactivation assay. This involves inserting a transcription factor's known binding site into a reporter vector, co-transfecting the reporter and an expression plasmid & gene of interest, then recording if the transcription factor activates or represses the reporter.

One possible method that could be used to analyse the function or otherwise of the potential splice site mutation would be to perform real-time PCR on tissues of interest. However, the low level of conservation of this base between splice sites, despite its proximity to the intron/exon boundary – would suggest that this mutation was less likely than the two coding mutations to cause a visible phenotype.

Chapter 5: Phenotyping of Mouse Mutant Foxf2^{W174R}

5.1 Introduction

In order to maximise the information from mutations that are identified with the gene-driven screen approach, mutant mice can undergo a systematic phenotyping procedure to detect subtle behavioural, morphological and metabolic deviations from the wildtype norm. The broad-based phenotyping platform SHIRPA [57] has previously been utilised in the preliminary characterisation of large numbers of ENU mutagenised mice, including the F1 male offspring of mutagenised animals that constitute the Harwell ENU archives [55].

The Foxf2^{W174R} founder male would have already been screened for abnormal phenotypes during this dominant mutation screen because all the individuals which contributed DNA to the archive were screened. Further screening of heterozygotes is therefore likely to produce results within the normal range. However, subtle phenotypic differences are more likely to be detected when comparing cohorts of mice that carry a known mutation with their wildtype littermates than when ENU mutagenised F1 mice are compared which each other. Therefore the SHIRPA screen remained a valuable tool despite the fact that the F1 founder male had already been subjected to many of these tests.

One other limitation of the original SHIRPA screen is that it was only likely to have detected dominant mutations because it is unlikely that one locus would be mutagenised by ENU twice in the same individual. The analysis of recessive phenotypes is critical for the effective characterisation of genes. Organs and tissues which express the gene of interest need to be the focus of phenotypic analysis because these are the most likely regions to express such a phenotype. Foxf2 is expressed in the respiratory, urinary and digestive tracts in common with its closest paralogue Foxf1, and is also expressed in tissues that don't express Foxf1 such as the sensory organs, limb buds, and CNS [188] including the brain. Those tissues where Foxf2 but not Foxf1 is expressed are of particular interest because the probability of redundancy is reduced.

Eye analysis: Foxf2 is expressed in the developing eye and the eye is one of only two tissues showing significant levels of Foxf2 in the adult [188]. The eye defects that were observed in Del36 mice [21, 82], highlighted the potential value of investigating genes from the deleted region that are expressed in the eye.

Mutations in Foxc1, the closest physical neighbour of Foxf2 and a closely related paralogue, are responsible for many cases of anterior segment dysgenesis in mice [81] and mutations in other Fox genes including FOXE3, FOXL2 and FOXC2 cause a wide variety of eye defects. The eye is recognised as sensitive to the dosage alterations of many genes because of the delicately balanced interacting pathways involved in eye formation [193]. The discovery of a Foxf2 mutation in glaucoma

patients by Vieira and Gregory-Evans [155] provided particular impetus for the detailed analysis of the eye in Foxf2^{W174R} mutants.

CNS: The central nervous system is another site in which Foxf2, but not Foxf1, is expressed during development [188]. A large proportion of the tests that constitute the SHIRPA platform are for behavioural phenotyping, and behavioural anomalies are an important indicator of CNS defects. SHIRA can therefore be considered as a primary screen for CNS phenotypes

Heterozygous carriers of each mutation that was identified in the lethal recessive screen discussed in chapter 3, underwent SHIRPA testing. Five of ten tested lineages exhibited distinct neurobehavioural phenotypes, three of which were associated with overt neurodevelopmental defects [26].

Limbs: There are high levels of Foxf2 but undetectable levels of Foxf1 in the developing limbs of mice [188]. SHIRPA includes a general appearance screen with a digit count and a grip strength test. These are both important primary screening methods for the detection of limb deformities.

Digestive tract: Following the publication of the Foxf2 knockout which described the presence of a cleft palate [153] in that mutant, it was considered important to determine whether the palate is affected in the Foxf2^{W174R} mutant.

Although not analysed during this work, the gut and tongue are also of interest. Malformation of the tongue was described by Wang in the original knockout publication [153]. Aitola noted differences in expression between Foxf1 and Foxf2 in the developing gut [188] and subsequently, Ormestad revealed a loss of structural integrity of the intestines and a wide spectrum of other gut phenotypes in Foxf2 homozygous knockouts and to a lesser extent in Foxf1/Foxf2 compound heterozygotes [154].

Lung analysis: As in the eye, expression of Foxf2 occurs in the adult as well as the developing lung. Foxf1 shows similar expression in the lung [188], suggesting possible redundancy. However, the continuing expression of the gene in the adult tissue warrants further investigation.

There is a very high probability that multiple mutations will be present in the genome of any individual that is subjected to ENU mutagenesis. These additional mutations could confound the characterisation of phenodeviants because abnormal phenotypes may result from combinatorial effects of multiple mutations. In order to reduce the probability of the presence of additional mutations, $Foxf2^{W174R}$ carriers were outcrossed for successive generations to non mutagenised mice. Mutant Phenotyping was delayed until the colony reached generation G4.

5.2 Heterozygote Screening

The first Foxf2^{W174R} heterozygous mutants were superficially normal at birth and SHIRPA was therefore considered an ideal preliminary screening platform from which subtle phenotypes could be uncovered.

A G4 cohort of 16 Foxf2^{W174R} heterozygotes and 16 wildtype littermate controls were selected to undergo testing in order to determine whether there was any difference in appearance, behaviour or metabolism. Due to the potential for gender-specific effects in some of the tests, males and females were equally represented within each genotype group (8 individuals of each sex, within each genotype group of 16). All of the mice were between 5-8 weeks old during the test period in order to minimise age related effects. Urinalysis and Intra-peritoneal glucose tolerance tests (IPGTT) were performed by the phenotyping core at MRC Harwell; SHIRPA behavioural tests were performed by the phenotyping core and the author. Startle and PPI of the initial cohort of 16 test-mice were also performed by the phenotyping core but subsequent Startle and PPI tests were performed by the author. Metabolic bleeds were performed by the phenotyping core and metabolic analysis of these samples was carried out by Dr T. Hough.

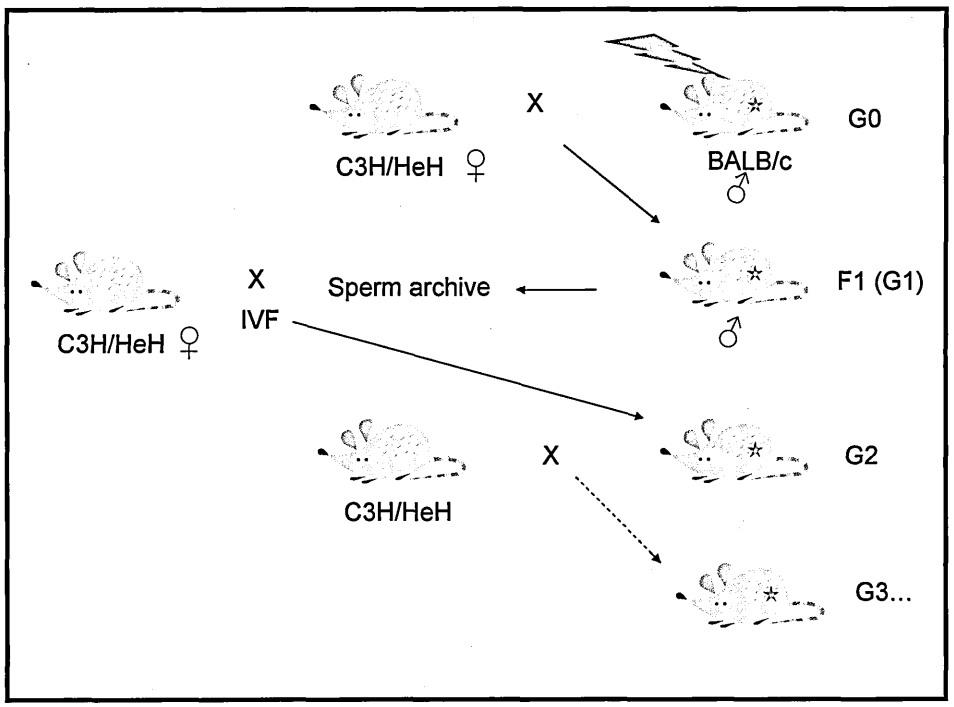


Figure 5.1: Gene driven screen breeding scheme
Mating scheme for mutant mice derived from the Gene driven screen. BALB/c males are mutagenised with ENU and mated with C3H/HeH females. Archived sperm from the F1 progeny of this mating is used for *in vitro* fertilisation with a C3H/HeH female to revive required mutant lines at the G2 generation. Several subsequent matings to C3H/HeH mice are recommended before analysing a mutant line, in order to reduce the probability of the presence of other confounding mutations.

5.3 Appearance

Systematic recording of the presence or absence of abnormalities in specific physical features was used to produce a detailed analysis of the appearance of the test mice. Traits which were examined were: Coat, Skin, Ears, Eyes, Teeth, Vibrissae, Limbs, Paws, Nails, Tail, Head shape and digits. Of these only the "Head shape" category resulted in any abnormal scoring. The heads of these animals were subtly shorter than their littermates; they were described as having a "short snout". It is likely that this was the result of a background strain effect however, as two of the abnormal heads belonged to wildtype individuals. Four heads were categorised as abnormal

amongst the mutants – all six abnormal heads from the cohort of 32 individuals, belonged to males.

The Harwell strain of C3H mice (C3H/HeH) have been known to produce offspring with subtly shortened heads (S. Polley, Somerset, R. Hardisty and H. Tateossian - personal communication) without any other noticeable abnormalities. There was no correlation between the mutation and the abnormal head shape – so this was not investigated further.

5.4 Behaviour

5.4.1 Behavioural Observations and Response to Simple Stimuli

Following the appearance observations, all mice undergoing SHIRPA are subjected to a range of behavioural tests. Initial testing is based on an assessment of undisturbed mice within a viewing jar. Gait, tail elevation, transfer arousal and body position are all monitored for abnormalities – none were seen in any of the tested animals. Unusual behaviour which can be characteristic of neurological or physiological defects is also looked for at this stage. Such activities could include head shaking or bobbing, circling, retropulsion (walking backwards), excessive grooming or spatial disorientation. None of these behaviours or any activity which appeared outside the normal behavioural range were observed in any individuals within the $Foxf2^{W174R}$ cohort.

Activity levels and levels of defecation were quantified by the number of grid squares entered within the arena and the number of faecal pellets produced. A student T-test was performed on each of these quantifiable behaviours. No significant difference between mutants and wildtypes was detected (See table 5.3).

One trait at this stage of screening which showed some variability between individuals was the transfer arousal observation. This is a test of how long the mice take to begin exploring upon entering the test arena. One wildtype male and one heterozygous female froze for longer than 5 seconds before moving. A larger number of the animals froze only briefly before moving; these included 6 male heterozygotes, 5 male wildtypes, 4 female heterozygotes and 5 female wildtypes. The remaining animals all moved immediately once they were within the arena. There did not appear to be any correlation between the mutation and any of these different behaviours.

The mice are exposed to an auditory stimulus – a click-box, and a touch stimulus – a metal rod lightly touching their nose; in order to observe their response. All animals responded to touch by the expected escape reflex. On hearing the click box, three of the eight heterozygote males tested did not exhibit the preyer reflex which is a characteristic sharp movement of the ear pinna in response to sound. None of the wildtype males lacked this response, although two wildtype and two heterozygous females also failed to exhibit the preyer reflex. The lack of this response could indicate hearing problems in these individuals. This trait was analysed in more detail

during the acoustic startle stages of the modified SHIRPA screen (section 5.5). All of the test mice had startle responses within the normal range, suggesting that the lack of preyer response did not indicate impaired hearing in these individuals.

Further spontaneous behavioural responses to different situations and stimuli were observed outside the arena. These tests were designed to test the motor reflexes, balance, coordination, and aggression of the animals. Observations included whether the mice struggled when either held by the tail or when scruffed; if the mice responded with a trunk curl when suspended by the tail; their ability to hang on to a wire suspended 20cm above their cage; pinnal reflex; corneal reflex; contact righting reflex; and whether there was any aggression or vocalisation from the mouse during the screening process.

Animal Group Size of Foxf2 W174R Mice Exhibiting Named Behaviours					
Test	Behaviour Exhibited	Male		Female	
		m/+	+/+	m/+	+/+
Body position	inactive	0	0	0	0
	Acitve	8	8	8	8
	Excessively active	0	0	0	0
Tremor	Present	0	0	0	0
Transfer arousal	Freeze for > 5 seconds	0	1	1	0
	Freeze briefly	6	5	4	5
	move immediately	2	2	4	3
Gait	Normal	8	8	8	8
	Abnormal	0	0	0	0
Tail elevation	Dragging	0	0	0	0
	Horizontal extension	8	8	8	8
	Elevated/straub tail	0	0	0	0
Startle response	None	0	0	0	0
	Preyer relex	5	8	6	6
	body move & preyer reflex	3	0	2	2
Touch escape	No reponse	0	0	0	0
	Response to touch	8	8	8	8
	Flees prior to touch	0	0	0	0
Additional Information	Head flicking/shaking	0	0	0	0
	Head bobbing	0	0	0	0
	Circling (intermittent)	0	0	0	0
	Circling (continuous)	0	0	0	0
	Retropulsion	0	0	0	0
	Spatial disorientation	0	0	0	0
Positional passivity	struggles when held by tail	3	1	3	1
	struggles when scruffed	8	6	7	6
Trunk curl	Present	0	0	0	0
Wire manoeuvre	Hangs on	8	8	8	8
	Falls off immediately	0	1	0	0
Limb Grasping	Present	0	1	0	0
Pinna reflex	present	8	8	8	8
Corneal reflex	Present	8	8	8	8
Contact righting reflex	Present	7	5	6	7
Biting or aggression	Evident during screen?	0	0	0	0
Vocalisation	Evident during screen?	0	0	0	0
Totals tested		**8**	**8**	**8**	**8**

Table 5.1: Behavioural testing
Animal group size of Foxf2^{W174R} heterozygotes (m/+) and their wildtype littermates (+/+) for the non-numeric behavioural tests within and above the arena and in the viewing jar.

Animal Group Size of Normal Physical Traits In Foxf2 W174R mice				
Physical Trait	Male	Male	Female	Female
	m/+	+/+	m/+	+/+
Coat	8	8	8	8
Skin	8	8	8	8
Head shape	4	6	8	8
Ears	8	8	8	8
Eyes	8	8	8	8
Teeth	8	8	8	8
Vibrissae	8	8	8	8
Limbs	8	8	8	8
Paws	8	8	8	8
Digits	6	8	8	8
Nails	8	8	8	8
Tail	8	8	8	8
Totals Tested	8	8	8	8

Table 5.2: Physical Traits
Animal group sizes of Foxf2^{W174R} heterozygotes (m/+) and their wildtype littermates (+/+) for all physical traits analysed prior to SHIRPA testing.

Foxf2 W174R Phenotyping Screen: Statistics						
Screen	**Test**	**Genotype**	**Sex**	**Mean**	**SD**	**P=**
Weight (g)	Weight (g)	m/+	M	25.74	2.17	0.65
		+/+	M	26.22	2.00	
		m/+	F	21.04	1.57	0.57
		+/+	F	21.43	1.07	
SHIRPA	Faecal pellets (number)	m/+	M	3.25	1.75	0.33
		+/+	M	4	1.2	
		m/+	F	4.5	2.67	0.08
		+/+	F	2.38	1.77	
	Locomotor activity (number of squares)	m/+	M	2.875	2.48	0.7
		+/+	M	3.38	2.72	
		m/+	F	4.5	2.39	0.34
		+/+	F	3	2.71	
Grip Strength	Forelimbs (arbitrary units of measuring equipment)	m/+	M	109.4	14.8	0.95
		+/+	M	108.7	22.2	
		m/+	F	96.7	17.5	0.78
		+/+	F	98.9	11.8	
	Fore and Hind limbs	m/+	M	207.9	13	0.15
		+/+	M	193.5	23.5	
		m/+	F	150.5	11.3	0.07
		+/+	F	161.7	11.4	
Spontaneous alternation	Latency (seconds)	m/+	M	42.63	50.6	0.39
		+/+	M	78	99.8	
		m/+	F	65	96	0.44
		+/+	F	37.18	41.3	
	% alternation (%)	m/+	M	47	32	0.5
		+/+	M	61	50	
		m/+	F	46	23	0.98
		+/+	F	46	28	
PPI	% inhibition 12Hz	m/+	M	67.5	10.2	0.009
		+/+	M	53.11	7.4	
		m/+	F	59.6	13.2	0.29
		+/+	F	50	16.4	
	% inhibition 20Hz	m/+	M	65.1	14.4	0.07
		+/+	M	53.2	17.1	
		m/+	F	51.1	17.6	0.91
		+/+	F	51	16.8	
Startle	white noise	m/+	M	1048.7	343	0.85
		+/+	M	1084.2	367.3	
		m/+	F	666.3	130.9	0.71
		+/+	F	632.7	217.9	
	response to 12Hz	m/+	M	1003.6	327.9	0.61
		+/+	M	1087	284.3	
		m/+	F	617	251	0.5
		+/+	F	697.8	217.7	
	response to 20Hz	m/+	M	1102.1	184.9	0.46
		+/+	M	1227.3	421.7	
		m/+	F	872.2	259	0.1
		+/+	F	676.9	188.5	

Table 4.3: Primary screen Statistics
A statistical summary of the behavioural phenotyping of Foxf2^{W174R} heterozygotes (m/+) and their wildtype littermates (+/+). Of those tests which could be compared statistically by the student T-test, the only test which indicated a significant difference between wildtypes and Foxf2^{W174R} heterozygotes was the PPI of males at 12 kHz. The single significant result, where $p \leq 0.05$ is highlighted in yellow.

In both gender groups, three mutants from a total of eight struggled when held by the tail, whereas only one wildtype in eight struggled. During the scruff test, only two wildtype males, one heterozygous female and two wildtype females did not struggle when scruffed.

Only one animal, a wildtype male, fell from the wire immediately on being placed there. This same animal was also the only individual which demonstrated the limb grasping reflex.

During the contact righting reflex test, the animal is placed in a Perspex tube and inverted. Several animals were slow to regain an upright position. These included one mutant and three wildtype males, and two mutant and one wildtype female. The remaining animals showed a normal righting reflex.

None of these tests showed a strong enough correlation between mutation and behaviour to be of interest. All other tests showed no variability between mutants and wildtypes in their normal response.

5.4.2 Light Dark Box Test

The Light-dark box automatically monitors the movement of test mice within an area containing both dark and light regions. They are a useful tool in the detection of a range of different mouse models for anxiety [249], which tend to spend a reduced time in the light areas. The parameters measured for this screen were: Latency, number of moves, time spent moving, velocity, distance covered, time spent in the

light area, time spent in the dark area, number of times the light area was entered, vertical plane entries (rearing), central area entries in the light region, central area entries in the dark region.

None of these parameters showed a significant difference between the wildtype and mutant heterozygotes when the light-dark box output was analysed with a student-t test (see figure 5.1 and appendix 5.8).

Foxf2 W174R Phenotyping Screen:

Statistics - (Light-Dark Box)

Test	Genotype	Sex	Mean	SD	P=
Latency (seconds)	m/+	M	20.7	14.9	0.87
	+/+	M	19.4	13.6	
	m/+	F	25.8	12.6	0.74
	+/+	F	29.5	20.8	
Moves (number)	m/+	M	35.3	9.4	0.19
	+/+	M	30.6	2.2	
	m/+	F	35.2	4.2	0.42
	+/+	F	38	6.3	
Move time (seconds)	m/+	M	250.5	10.9	0.14
	+/+	M	257.5	5.5	
	m/+	F	254.6	9.8	0.49
	+/+	F	250.2	10.5	
Velocity	m/+	M	799.9	161.5	0.29
	+/+	M	869.4	69.1	
	m/+	F	905.4	107.6	0.36
	+/+	F	819.7	145	
Distance (cm)	m/+	M	159.9	32.3	0.46
	+/+	M	173.9	14	
	m/+	F	181	21.5	0.3
	+/+	F	163.9	29	
Time in the dark (seconds)	m/+	M	132.2	22	0.47
	+/+	M	141.3	22.7	
	m/+	F	139.8	28.9	0.8
	+/+	F	145.2	37.8	
Time in the light (seconds)	m/+	M	167.8	22	0.47
	+/+	M	158.8	22.7	
	m/+	F	160.2	28.9	0.8
	+/+	F	154.9	37.8	
Entry into the light	m/+	M	10	3.5	0.72
	+/+	M	10	2	
	m/+	F	11	2.3	0.29
	+/+	F	12.5	2.1	
Rearing	m/+	M	49.2	10.4	0.6
	+/+	M	46.3	9.6	
	m/+	F	43.4	5.2	0.94
	+/+	F	43.8	11.6	

Table 5.4: Light-Dark box testing
This test is based on the premise that there is a direct correlation between increased exploratory behaviour and reduced anxiety [250]. There were no significant differences between Foxf2^{W174R} heterozygotes (m/+) and their wildtype littermates (+/+) in any of the tested parameters.

5.4.3 Spontaneous Alternation

The spontaneous alternation test involves placing a mouse at one end of a Y-shaped arrangement of Perspex tubes, in which the mouse is free to explore. This is a test designed to detect defects in spatial learning. The latency period before the mouse leaves the starting arm and the number of alternations between arms are recorded.

There was no significant difference between wildtypes and mutant heterozygotes in either latency or alternation percentage on analysis with a 2-tailed student T test.

5.4.4 Swim Test

Mice are lowered into water at 37°C, and allowed to free swim for 1 minute. Any unusual swimming patterns are recorded. The swim test is one method of detecting vestibular dysfunction in mice.

No unusual swimming patterns or immobility was detected in any of the test mice.

5.4.5 Neuromuscular test (Grip Strength test)

Forelimb and hindlimb strength is measured during this test. There were no statistically significant differences between the mutants and wildtypes on analysis of the results with a 2-tailed student T test.

5.5 Startle Tests

A brief, strong acoustic stimulus can induce a "startle" reflex in vertebrates – a protective behavioural reaction consisting of an involuntary motor response. Tests for acoustic startle can provide evidence of deafness, or potentially for evidence of defects in the fear-potentiated startle reflex. This is a spontaneous reaction associated with the basolateral amygdala region of the brain, via the thalamus and perirhinal or insular cortex[251]. Abnormalities in this neural pathway are thought to be responsible for several anxiety disorders [252], and have also been identified in patients suffering from the early stages of Parkinson's disease[253].

No differences between the wildtypes and mutant heterozygotes were seen on analysis of the startle output from the original cohort with a 2-tailed student T test (see table 5.3)

T-test comparisons were also made between those individuals that did not exhibit a preyer reflex on exposure to click box stimulation with wildtypes exhibiting a preyer reflex. Startle data from the male preyer deficient individuals were compared with each of the wildtype males ($p=0.79$(white noise); $p=0.97$ (12 kHz); $p=0.94$ (20 kHz)). Because the female preyer deficient individuals were heterozygotes and wildtypes, two different comparisons were made: All preyer deficient females were compared with the preyer exhibiting wildtypes in the first analysis ($p=0.86$ for white noise; $p=0.54$ for 12 kHz, $p=0.25$ for 20kHz.). This was followed by a comparison

of only the heterozygous preyer deficient females with the preyer exhibiting wildtypes ($p=0.37$ for white noise; $p=0.28$ for 12 kHz, $p=0.28$ for 20kHz.).

Despite the absence of a preyer reflex in some individuals, they seemed to respond to acoustic startle within the normal range demonstrated by the preyer exhibiting wildtypes. Absence of the preyer reflex is thought to be an indication of severe hearing deficiencies – it is not thought to be a useful test for mild deafness. It can therefore be concluded that the absence of a preyer reflex in these individuals was not an indication of deafness.

Further Startle experiments were carried out to complement the additional PPI tests which were required to investigate the difference in PPI between mutant and wildtype males using a 12 kHz prepulse (See section 5.6).

The acoustic startle programme Startle-1 was used to test a further 7 males (m/+ (n) =3; +/+ (n) =4) which were used to supplement the original dataset. No significant difference between mutants and wildtypes was seen with this dataset either by itself ($p=0.70$ (white noise), $p=0.96$ (12 kHz), $p=0.91$(20 kHz)), or in combination with data from the original cohort.

5.5.1 Problems with Programmes Startle-1 and PPI-1

The sound waves produced at either 12 kHz or 20 kHz in either PPI or acoustic startle testing, should increase gradually to a plateau then gradually reduce towards

the end of each pulse. However, when measurements were made of the 12khz and 20khz sound bursts, the waveforms were spiking at two different time points within each test pulse (A. Parker, E. Coghill and F. Mackenzie: personal communication).

PPI and startle programmes which were used in the initial screening, PPI-1 and Startle-1, both contained 10ms pulses at 12 kHz and 20 kHz. The information from the waveform measurements therefore suggested that data from the original tests at these frequencies may not be reliable. Further PPI and acoustic startle tests were therefore required with programmes PPI-2 and Startle-2

5.5.2 Startle-2

A further two cohorts were screened, but with an alternative set of parameters – using the programme Startle-2. The first batch to be screened with Startle-2 (cohort 2) contained both males (m/+ (n= 4); +/+ (n=9)) and females (m/+ (n=12); +/+ (n=12)). The low number of heterozygous males in this batch meant that further screening was required. However, this cohort was still analysed individually by two-tailed T-Test. Neither the females (p=0.46) nor the males (p=0.79), showed a significant difference between mutants and wildtypes in their response to the white noise frequency, which is the only frequency used in Startle-2. as described in section 2.2.28

In the second of these cohorts (cohort 3), which consisted of males only (m/+ (n=8); +/+ (n=10)), a significant difference was observed between the mutants and

wildtypes (p=0.012) (See figure 5.2). These data were combined with the other Startle-2 dataset and analysed with a 2-tailed T-test. A significant difference between heterozygote and wildtype males was not seen in the combined Startle-2 data (p=0.175). The two datasets combined therefore do not show a significant difference between mutants and wildtypes. However, various factors such as the increased environmental control, larger cohort size and reduction in background heterogeneity meant that the second tests were more reliable. Therefore there is a good chance that there is a genuine startle response in these mice.

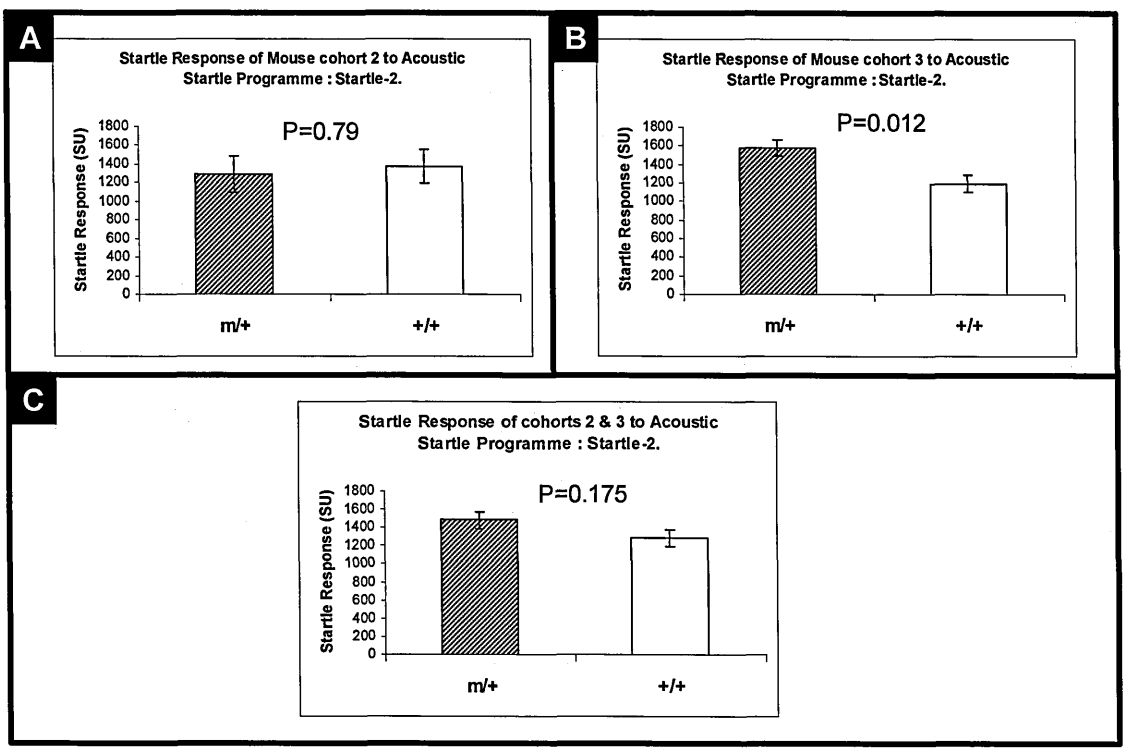


Figure 5.2: Startle Response to Startle-2
Response of 2 cohorts of Foxf2^{W174R} heterozygotes (m/+) and wildtype littermates (+/+), which were tested with Startle-2. The mutants demonstrated a significant startle increase in one cohort but the other showed an insignificant decrease. The combined data did not show a significant difference between mutant and wildtype.

5.5.3 Analysis of the Startle Data

Although there was no significant difference in the acoustic startle data from initial tests using the Startle-2 programme, the significant difference seen in the larger of the two cohorts (cohort 3) warranted further analysis.

These two cohorts were from different generations and had been raised in different environments. Cohort 2 which were from generation G5, were raised and tested in a standard mouse facility (building 371 of MRC Harwell). Cohort 3 were G8 and were raised in a facility with individually ventilated cages (the Mary Lyon Centre), then transferred to building 383 before testing. It is possible that either these differing environments or the distance in generations from the founder male could have affected the animals' response to the startle test. Previous experiments have demonstrated that subtle alterations in environmental enrichment conditions can affect the response of mice to various forms of behavioural phenotyping assessment [67]. A generational effect could have been caused by a loss of additional ENU induced mutations, or by a reduction in the BALB/c to C3H/HeH background ratio.

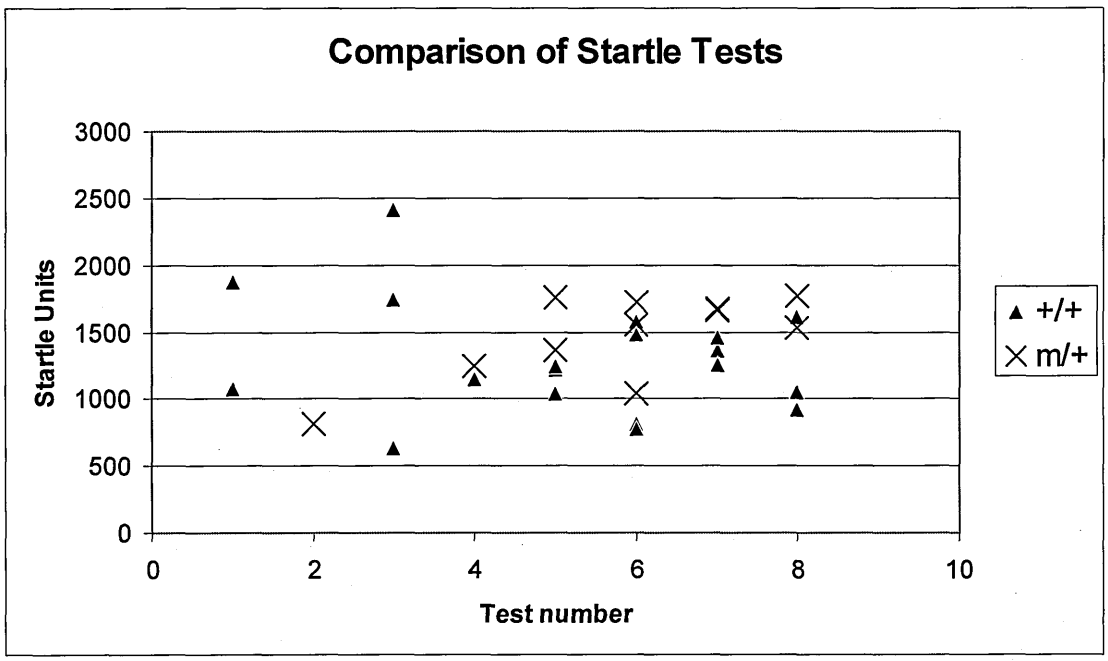


Figure 5.3 Startle Data for Startle-2 Programme
A comparison of Startle data from all individuals tested with Startle-2. Test numbers 1-5 were tests performed on cohort 2 in building 371, test numbers 6-8 were performed on cohort 3, which were mice that were born and raised in the Mary Lyon Centre and tested in building 383. Wildtype (+/+) startle responses are represented by a triangle and Foxf2W174R heterozygote responses are represented by a cross.

The order of testing was compared with the startle data to compare data ranges between cohorts 2 and 3. The data obtained from cohort 2 (tests1-5 of figure 5.3) did seem to be more widely dispersed than in cohort 3 (tests 6-8). This could be viewed as evidence that there is an effect on the data caused by the different test conditions between the two sample groups. A more diverse range of startle responses could also be an indication of the less genetically homogenous population in cohort 2 which were G5 rather than G8 mice.

There are known to be differences between in-bred mouse strains in their response to acoustic startle [254]. BALB/cJ and C3H/HeJ were amongst the 40 strains compared

for acoustic startle response in a study by James Willott and Kenneth Johnson included in the Mouse Phenome Database [255]. They identified a large and significant increase in the startle response of BALB/cJ individuals by comparison with C3H/HeJ mice (a mean difference in males from ~120SU at 70dB to ~600SU at 100dB). This may explain the wider range of data points in earlier generations that may have masked the difference in the earlier data (figure 5.3). Cohort 2 was from generation G5 so these mice would carry a higher percentage of BALB/c derived genome than cohort 3, which were G8 mice.

Equally though, the elevated startle in BALB/c mice allows the possibility that heterozygous mice may exhibit higher startle responses due to a BALB/c allele with close linkage to Foxf2.

When histograms were constructed to view the data distribution for the wildtypes and heterozygotes within the combined dataset, there looked to be two clearly different response ranges for the two genotypes, despite possible differences caused by the sample conditions (see figure5.4). Heterozygotes were more numerous at the 1625-1750 SU range and wildtype mice were more frequent at the 1000-1125 SU range. There were definite slopes leading to these values with adjacent ranges also showing high frequencies, adding weight to the hypothesis that the heterozygotes and wildtypes have different startle responses.

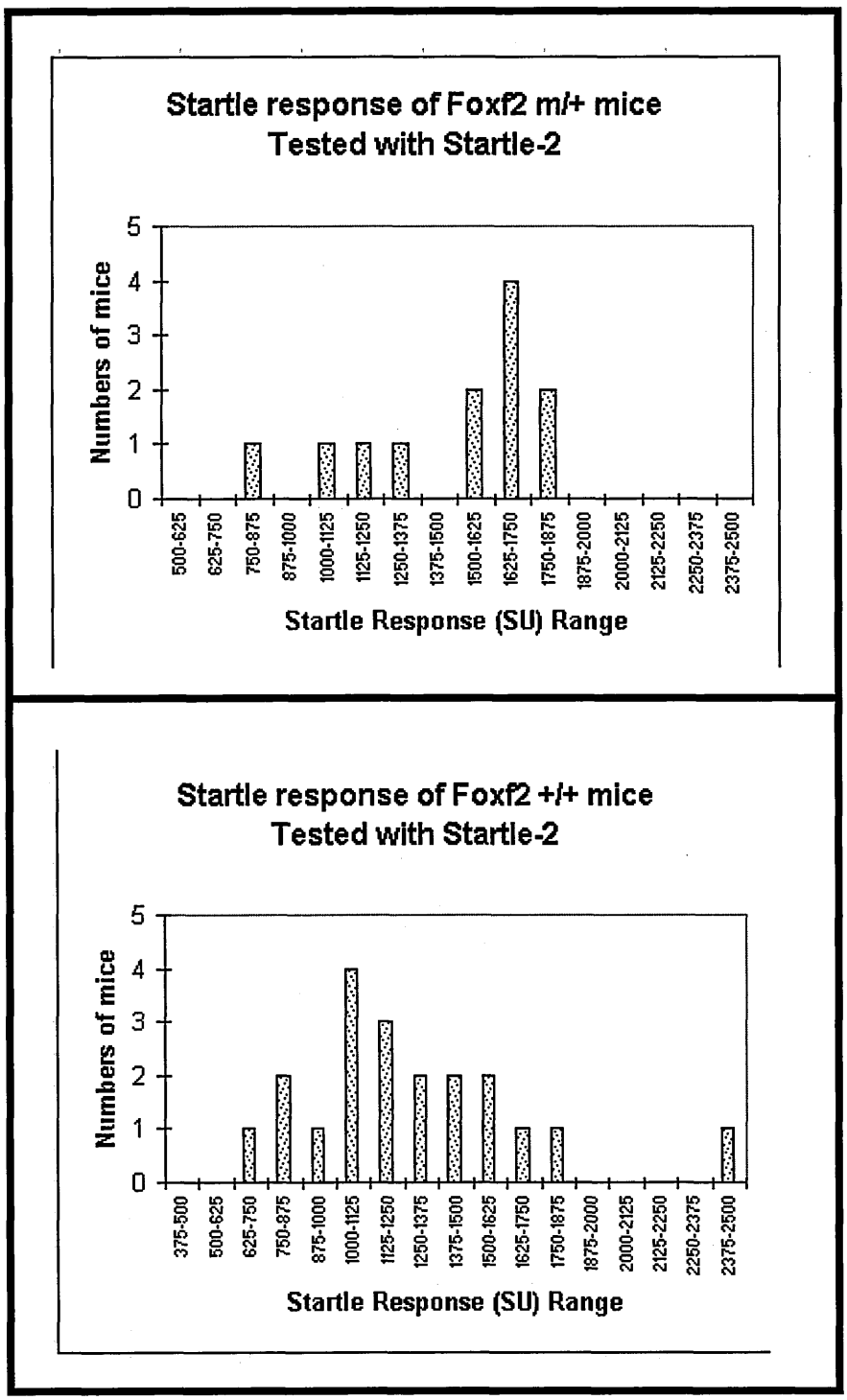


Figure 5.4: Histograms for Startle-2 Response
The numbers of mice at a range of startle response (SU). The heterozygous mice (m/+) have a greater frequency at the 1625-1750 SU range, whereas the greatest frequency for the wildtype mice (+/+) is at the 1000-1125 SU range.

One of the results from the wildtype data looked to be an outlier when viewed on both the histogram and test comparison plot. A box and whisker plot was therefore constructed to determine whether this was the case (figure 5.5). There is a general rule in statistics which states that if a data point is below 1.5 X the interquartile range below the first quartile, or 1.5 X the interquartile range above the third quartile, it can be viewed as being too far from the central values to be reasonably included.

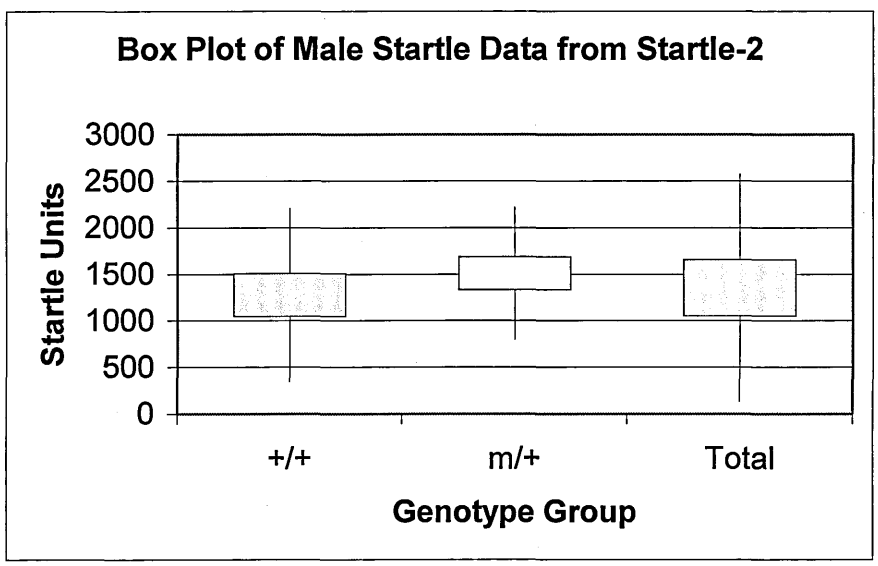


Figure 5.5: Box-plot of startle-2 data
Startle-2 data from wildtype (+/+) and Foxf2^{W174R} heterozygote (m/+) littermates. One value (2413 SU for individual 24.1e) from the wildtype mice could be considered an outlier because it is above Q_3 + 1.5×IQR (Interquartile Range) within the wildtype group. The boxes represent the interquartile range for each dataset and the whiskers represent below Q_1 – 1.5×IQR and above Q_3 + 1.5×IQR, beyond which, data could be considered outliers.

Individual 24.1e showed a startle response of 2413 SU. This was a very high startle response compared with all the other animals tested, and particularly when compared with other wildtypes. It was classified as an outlier within the wildtype

group ($Q_3 + 1.5\times IQR = 2204$) or the heterozygote group ($Q_3 + 1.5\times IQR = 2218$). However, when both groups were combined ($Q_3 + 1.5\times IQR = 2572$), it was not considered an outlier.

If individual 24.1e is classified as an outlier; the T-test that compares heterozygotes and wildtypes gives a p value of 0.04 when applied to the whole Startle-2 dataset, a significant, albeit a mildly significant result. There is no doubt that these are an inconclusive set of results. It cannot be concluded from this data that $Foxf2^{W174R}$ heterozygotes have an increased startle response in comparison with wildtypes, but nor can this hypothesis be discounted.

When viewed as individual cohorts however, the weight of evidence – with a highly significant p value of 0.012 in the larger, later cohort suggests that there is a difference induced by the mutation that had been previously been masked by other factors. None of the individuals scored below 500SU during testing, so deafness is unlikely to be a factor contributing to the difference.

The secondary startle tests were only performed because of the significant difference in PPI response between mouse cohorts at 12kHz. Acoustic startle testing is inherently linked to PPI tests and the two tests are often performed together in order to better define any phenotype that is highlighted. Although the significance of the PPI result is less convincing following subsequent PPI tests and the correction for multiple testing, it is possible that the significant results of the two tests are

phenotypically linked. Two significant results from such closely related tests could be construed as acquiring a degree of cumulative credibility. However, although both tests measure a startle response in the experimental subject, startle has been shown to be linked to the spinocerebellar region of the brain whereas PPI is linked to the midbrain and cortex regions.

Each of the regions of the CNS that were analysed by Aitola [188] expressed Foxf2 during embryonic development, but there was temporal variation as well as varying strengths of expression between the regions. The strongest signal appeared in the Entorhinal cortex at 36 hours after birth. Strong signals relative to other CNS regions were also seen in the Pia mater at between E17 and 36 hours after birth, the subicular area at E17 and at 3 days after birth and in the tectum between at E11 (the earliest stage measured) to E13.

5.6 Pre–pulse Inhibition (PPI)

Pre-pulse inhibition is the reduction of the startle response by exposure to a milder, non startle inducing stimulus immediately before the acoustic startle (10-500 milliseconds before the startle in mammals). It has proved to be a valuable tool for assessing deficits of sensorimotor gating and information processing in a broad range of psychiatric disorders such as schizophrenia, obsessive compulsive disorder and Tourette's syndrome (reviewed by Braff et al [256]).

Mice are placed onto a transducer platform within an acoustic chamber. The whole-body response of the mouse is measured in reaction to the stimulus in an acoustic startle test, in which only the startle inducing stimulus is produced. Once the baseline data for the responsiveness of the mouse is established, further acoustic tests which include both the startle pulse and pre-pulse can be performed and the inhibition percentage can be calculated at a range of frequencies

The initial screen involved a PPI programme (PPI-1) which subjected the mice to a white-noise startle, preceded by a pulse of either 12kHz or 20kHz. as described in section 2.2.28.

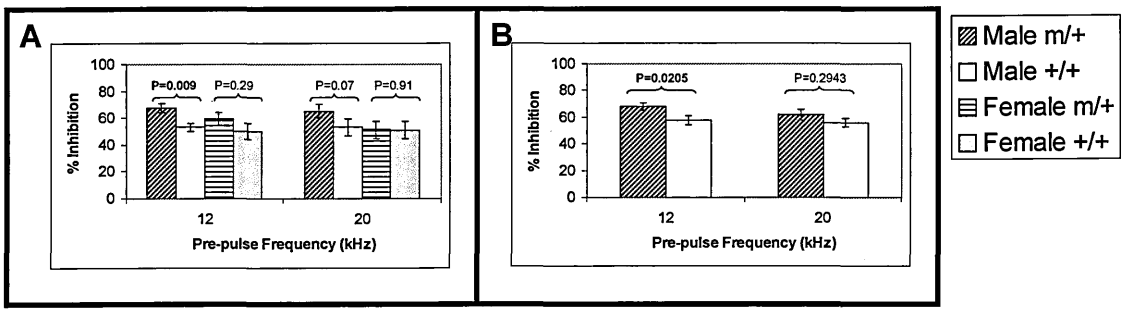


Figure 5.5: PPI-1 Response
Pre-pulse inhibition with programme PPI-1 of the initial cohort (A), and this data from males only with the addition of data from 7 further individual males (3 heterozygotes, 4 wildtypes) to the dataset (B). 2-tailed student T-tests suggested that there was a significant difference ($p<0.01$) between heterozygote (m/+) and wildtype (+/+) males when exposed to a pre-pulse of 12kHz in the original cohort (A). However, the small subset of males which were added had a marked effect on the p value which was increased to 0.02 following their inclusion in the analysis.

On analysis of the PPI startle data, 2-tailed T-tests gave a P value of 0.009 in comparisons between male heterozygotes and wildtypes when exposed to a prepulse of 12kHz (see figure 5.2a). There was no significant difference between male heterozygotes and wildtypes at 20kHz although again - there was an increase in the mutants, and at p=0.07, the result was close to the standard significance threshold of 0.05. No significant difference was seen between female heterozygotes and wildtypes at either 12kHz or 20kHz.

The males from original cohort of mice were retested with PPI programme PPI-2. This programme runs test pulses to determine the baseline response of the mice at 70dB, 80dB, 85dB and 90dB. Tests are then carried out in which a 110dB pulse is preceded by pre-pulses of 70, 80, 85 or 90dB which occur 50msec before the 110dB pulse. All tones were at the white noise frequency. Full details of the PPI parameters are described in section 2.2.28. There was no significant difference between heterozygote and wildtype males in these retests ($p > 0.05$ at 70dB, 80dB, 85dB and 90dB).

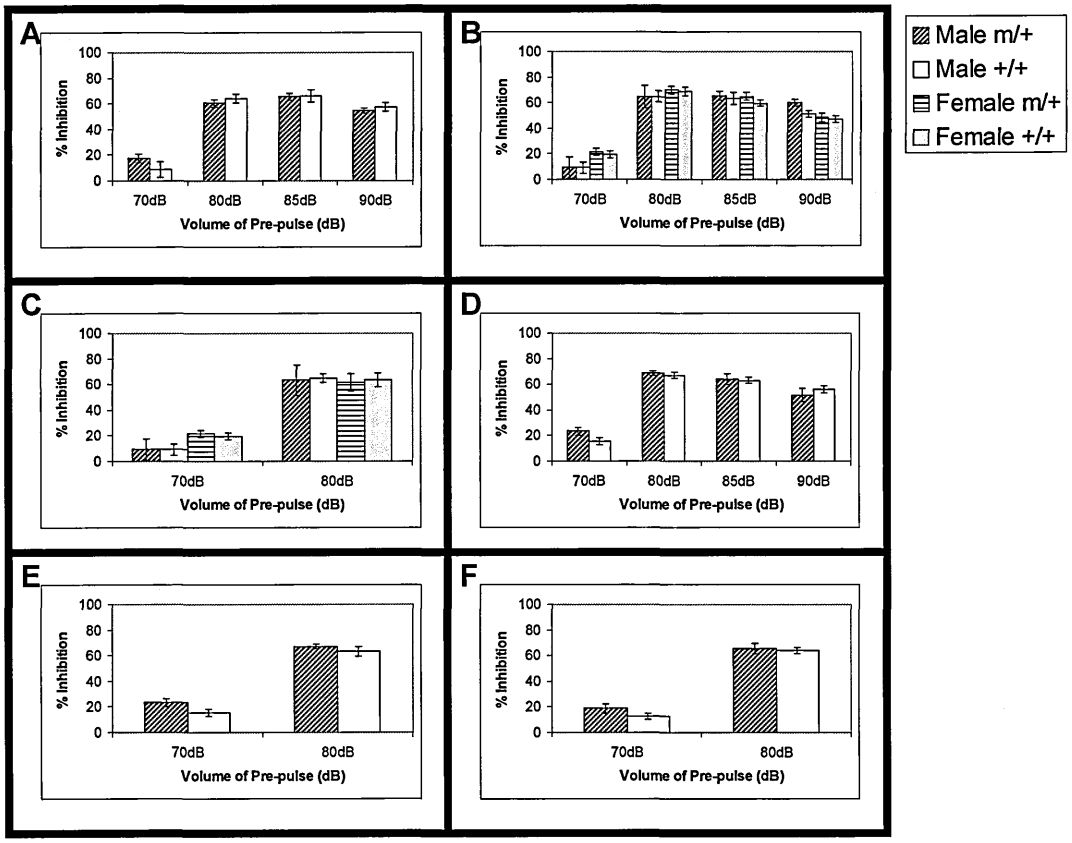


Figure 5.6: PPI-2 Response
The response of Foxf2^{W174R} heterozygote (+/m) and wildtype (+/+) mice to the programme PPI-2. Following tests with PPI-1, males from the original cohort were retested with PPI-2 (A). A second cohort of males (m/+(n)=4; +/+(n)=9) and females (m/+(n)=12; +/+(n)=12) at G5 was tested with the PPI-2 experiment (B). Startle data from prepulse and pulse baseline tests was used to exclude data from animals which demonstrated either a significant startle response (>350 SU) to pre-pulse volumes alone (70, 80, 85 or 90dB), or a very low startle response (<350SU) to pulse volumes of 110dB. This resulted in the exclusion of all PPI data where an 85dB or 90dB pre-pulse was used, all 70dB data remained unchanged, 80dB data had the following alterations to test size: +/-males(n)=3; +/+ males(n)=4; +/- females(n)=4; +/- females(n)=7 (C). A third cohort of mice at G8 which included only males (m/+(n)=8; +/+(n)=11), was tested for PPI in order to increase the amount of data for heterozygous mutant males (D). This data was modified by the exclusion of data which showed low startle response to a test pulse or a high startle response to a pre-pulse test as previously described. Again, no alterations were needed for the 70dB dataset and the 85dB and 90dB datasets were excluded. Data for the 85dB dataset contained the following alterations to test mice numbers: +/-males(n)=5; +/+ males(n)=6 (E). Both sets of modified data were combined (F). None of the data showed a significant difference between wildtypes and Foxf2^{W174R} heterozygotes.

Seven additional individual males were tested with the original PPI programme (PPI-1). 2-tailed student T tests for this small subset of animals resulted in a P value of 0.99. When combined with the original data set, the increase in PPI observed in

the heterozygotes remained significant, but the P value was increased to 0.0205, so the difference in this larger population was less significant than in the original cohort.

The reduction in pre-pulse inhibition is a trait that can be symptomatic of schizophrenia or bipolar disorder in humans. PPI was significantly higher in $Foxf2^{W174R}$ heterozygotes than in their wildtype littermates in preliminary tests with programme PPI-1. Screens for models of psychiatric disorders would tend to favour those animals with a lower than average PPI in order to mirror those conditions.

Despite this, it was considered worth investigating the effect further. An allele that results in a high PPI could provide an interesting insight into the activity of the gene and its potential role in the neural pathway involved in pre-pulse inhibition. A mutation at an alternative allele within such a gene could have a different effect on gene activity and may result in a reduction or an increase in PPI. A genuine increase in PPI could therefore highlight Foxf2 as a potential candidate for schizophrenia and related disorders. Therefore further investigation of the PPI of these mutants was considered important to assess the robustness of the preliminary results.

5.6.1 PPI-2: Retesting of the Initial Male Cohort

On retesting the first cohort with programme PPI-2, the difference between mutants and wildtypes was not significant. Previous studies have indicated that there may be a test-retest effect with PPI [257, 258] and preconditioning may have affected the response of the test mice and obscured the difference between the mutant and

wildtype individuals. This second dataset did, however, suggest that the difference seen in the initial dataset may not be as robust as initially thought.

5.6.2 Is there a difference in the PPI Response of Foxf2^{W174R} Mutants?

The significant difference between male heterozygotes and wildtypes at 12kHz, which was seen in the output from PPI programme PPI-1, looked less convincing following the addition of data from more individuals tested with the same method. However, most of the additional testing carried out used the PPI-2 programme, which did not incorporate a pulse at 12kHz. Further testing is required to completely rule out a PPI difference between mutants and wildtypes, but the combination of data from the additional small PPI-1 dataset, PPI-2 data and results from the PPI-1 retest build a strong case for the absence of a genuine difference.

5.7 Metabolism Tests

5.7.1 Blood Tests Using the Olympus AU400 Clinical Chemistry Analyser

The mice were tested for levels of a range of different blood biochemical markers as described in section 2.2.29. Testing occurred initially at 5 months on 05/08/04. Mice were then aged for an additional 3 months before retesting at 19/11/04 to detect later onset phenotypes.

The following parameters were measured: Sodium, potassium, chloride, urea, creatinine, total calcium, phosphate, ALP, ALT, AST, total protein, albumin, total

cholesterol, triglycerides and glucose. Unfortunately, not enough blood was collected for testing of the complete cohort so there was not a full dataset.

Triglycerides were not measured in a sufficient number of individuals for comparisons between wildtypes and heterozygotes of either sex. Total cholesterol was not measured in enough females for comparisons of heterozygous and wildtype females. Each of the tests were performed on varying numbers of samples from within each mouse sex/genotype group as described in table 5.1.

Those tests which were performed on at least 3 individuals from each of these groups were compared by T-test. Data from test 1 and test 2 were analysed individually where possible, and a comparison between heterozygotes and wildtypes of both sexes within the combined dataset was also made.

Results from the first test showed potassium, chloride and calcium to be significantly higher in male heterozygotes than male wildtypes, and urea was significantly higher in heterozygote females than wildtype females (figure5.8). The level of confidence in the significant difference between heterozygotes and wildtypes with the calcium (males) and urea (females) was at $p < 0.05$, whereas a greater level of confidence is seen in the potassium and chloride levels of males where $p < 0.01$. However, if the Bonferroni correction for multiple testing is applied as described in appendix 5.11, then the significance threshold for each of these parameters would fall well below 0.05 to 0.001 for females or 0.0008 for males –

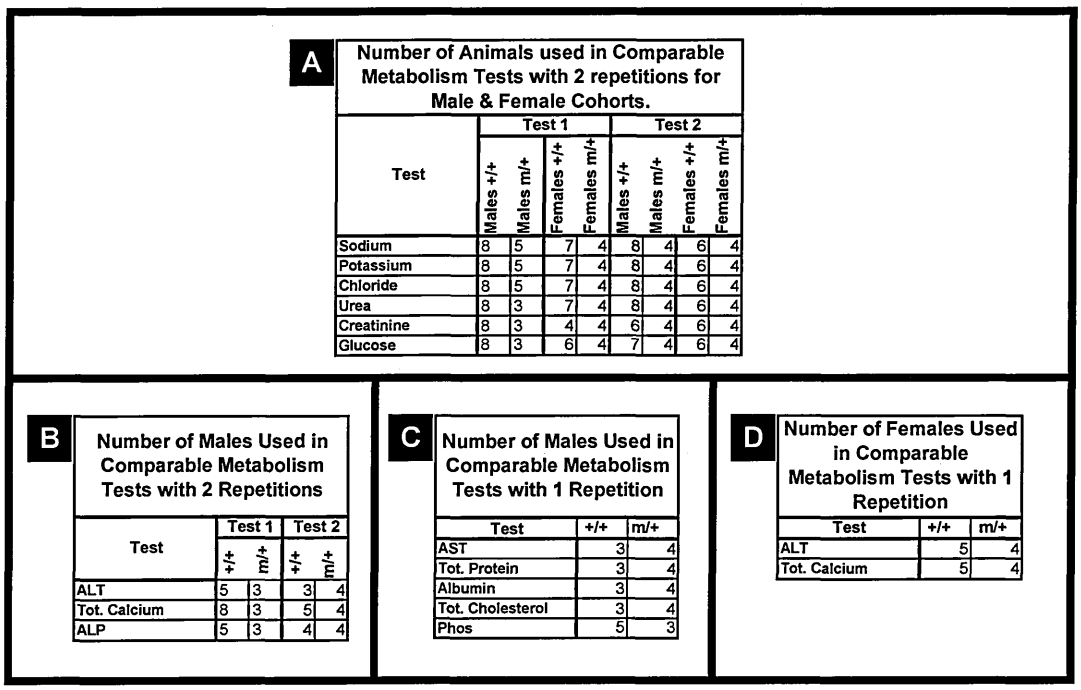
thus rendering the results insignificant. It is surprising however, that so many of the significant results that arose from the phenotyping screen, occurred during metabolic testing. With the high numbers of tests performed, type II errors are to be expected, but four significant results from the metabolic screen is unusually high with respect to the relatively small number of tests in the metabolic screen.

The results from the second set of tests, on the mice after they had aged for 3 months, showed no significant difference between heterozygotes and wildtypes in any of these tests.

Despite this though, the potassium result remains of particular interest in light of recent work by Philip-Couderc et al, who identified human FOXF2 as a regulator of ATP-dependent potassium channel (KATP) sub-units [231]. When HL-1 mouse atrial cells were transfected with a human FOXF2 plasmid, the KATP proteins KIR6.1, KIR6.2 and SUR2A were upregulated – and there was a tenfold increase in FOXO1/FOXO3 expression, which in turn were also shown to be potent activators of KATP sub-unit expression. A highly significant positive association with KIR6.1, SUR1A, SUR2A and SUR2B was seen in ventricular tissue following the induction of myocardial infarction.

Interestingly and in apparent contrast with these findings; FOXF2 showed a negative association with KIR6.1 and SUR1A in rat neonate cardiomyocytes. Despite the upregulation of FOXO genes by FOXF2, FOXO over-expression has an inhibitory

effect on the expression of FOXF2, indicating the genes involvement in a negative feedback system.



A Number of Animals used in Comparable Metabolism Tests with 2 repetitions for Male & Female Cohorts.

Test	Test 1				Test 2			
	Males +/+	Males m/+	Females +/+	Females m/+	Males +/+	Males m/+	Females +/+	Females m/+
Sodium	8	5	7	4	8	4	6	4
Potassium	8	5	7	4	8	4	6	4
Chloride	8	5	7	4	8	4	6	4
Urea	8	3	7	4	8	4	6	4
Creatinine	8	3	4	4	6	4	6	4
Glucose	8	3	6	4	7	4	6	4

B Number of Males Used in Comparable Metabolism Tests with 2 Repetitions

Test	Test 1		Test 2	
	+/+	m/+	+/+	m/+
ALT	5	3	3	4
Tot. Calcium	8	3	5	4
ALP	5	3	4	4

C Number of Males Used in Comparable Metabolism Tests with 1 Repetition

Test	+/+	m/+
AST	3	4
Tot. Protein	3	4
Albumin	3	4
Tot. Cholesterol	3	4
Phos	5	3

D Number of Females Used in Comparable Metabolism Tests with 1 Repetition

Test	+/+	m/+
ALT	5	4
Tot. Calcium	5	4

Table 5.5: Metabolic screen

Number of individuals used for each of the Metabolic measurements that were comparable by T-Test between Foxf2W174R heterozygote (m/+) and wildtype (+/+) cohorts. Some of the bleeds did not provide enough blood for measurements of all potential parameters. In cases where fewer than 3 individuals were successfully tested, a t-test has not been presented here, but the data is available in appendices 5.3-5.4. Image A shows those tests for which enough measurements for comparisons were obtained on two dates for both the male and female cohorts. B lists the tests where enough measurements were obtained on two dates for males only. C lists the tests where measurements from one date were comparable in males. D lists those tests for which measurements of females were obtained from one date only.

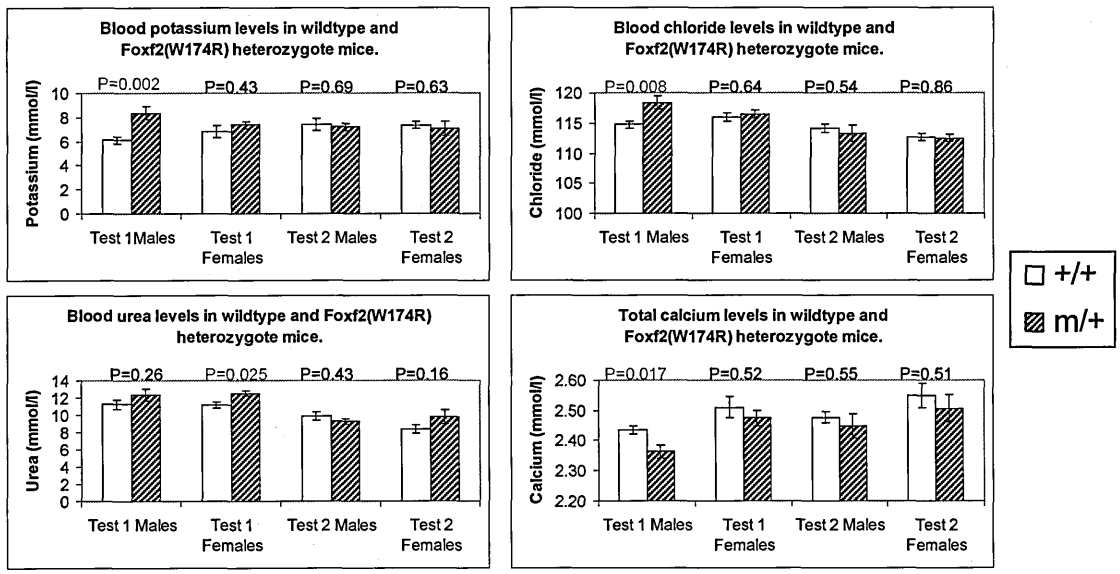


Figure 5.8: Significant results from metabolic screen
The blood biochemical markers which showed a significant difference between wildtypes (+/+) and heterozygotes (m/+). Data from each metabolite was compared using a 2-tailed student-T test. None of the parameters which showed a significant difference in test 1, showed a significant difference in test 2. Significant P values are shown in red.

5.7.2 Urine Dipstick Tests

Urine dipsticks were used for approximate measurements of the following metabolites/constituents in the urine of the mouse cohort: Bilirubin, urobilinogen, ketones, ascorbic acid, glucose, protein, blood, and nitrites, alongside pH. The numbers of individuals varied within each mouse group, there were 6 heterozygous males, 8 wildtype males, 5 heterozygous females and 6 wildtype females.

The negligible levels detected of Bilirubin, glucose, and blood showed no variation between any of the individuals. The urobilinogen test also showed no variation between individuals in their "normal" reading. On measuring ketones, one wildtype male individual was recorded as "negligible" whereas the other individuals were recorded as "+/neg". The dipstick documentation lists the lowest detectable levels of

acetoacetic acid as 10mg/dl and the "+" reading as approximately 25mg/dl, so a +/neg can be interpreted as between 10 and 25 mg/dl acetoacetic acid [259]. This is above the expected level of <5mg/dl acetoacetic acid levels that are expected in healthy mice [260]. The nitrite test indicated that one heterozygote and one wildtype female had "negligible" nitrite levels whereas all the other mice had "none".

It was immediately apparent that these tests showed very little variability. The slightly elevated ketone levels in all but one individual could be an indication of diabetes mellitus or a number of other metabolic disorders [261] and suggests that additional mutations may be present in the Foxf2^{W174R} line which have the potential to confound the phenotypic testing of metabolism.

It is possible however, that the dipsticks were incorrectly read for this test. Interpreting the colour information on urine dipsticks and inferring intermediate datapoints between the basic "Negligible", "+", "++" and "+++" categories is only an approximation and a technician who is inexperienced with the analysis of this type of data could have inadvertently elevated the results. If the ketone elevation was real and caused by another mutation, the expected numbers of Foxf2 wildtype individuals expressing the phenotype would be low, whether or not there was linkage to Foxf2 – because wildtype littermates were also screened. These ketone results are therefore likely to be false positives.

Protein levels and ascorbic acid levels both exhibited a higher level of variability, and further analyses of these datasets were required. The readings for levels of each of these constituents were indicated on the urine stick as either "neg", "+", "++", "+++" or intermediate values where the reading looked to fall between 2 of these defined parameters. The approximate amounts of each constituent represented by the symbols that were used are shown in table 5.6.

Symbol	Protein (mg/dl)	ascorbic acid (mg/dl)
neg	0	0
neg/+	15	10
+	30	20
+/++	65	30
++	100	40

Table 5.6 Protein and Ascorbic acid dipstick unit conversion
Key to urine dipstick measurements of the approximate amounts of protein and ascorbic acid.

Ascorbic Acid Concentrations in wildtype and W174R heterozygote mice				
	+/m		+/+	
Ascorbic Acid (Dipstick Reading)	Number of males	Number of females	Number of males	Number of females
neg				
neg/+				
+				
+/++		4	4	2
++	6	1	4	4

Protein Concentrations in wildtype and W174R heterozygote mice				
	+/m		+/+	
Protein (Dipstick Reading)	Number of males	Number of females	Number of males	Number of females
neg		1		
neg/+			5	
+	4	3	2	6
+/++	2	1	1	
++				

Table 5.7: Ascorbic Acid and Protein Frequencies
Frequencies of mice with the approximate concentrations of ascorbic acid and protein in their urine, as determined by the urine dipstick method.

Comparisons between different groups of frequency data are often made using chi-squared tests. However, this test is only approximate and is thought to be unreliable for small datasets, particularly when comparing ordinal data (data which is categorised and there is a logical order between the categories). A test that is more appropriate for the ascorbic acid and protein datasets (both of which are small and ordinal) is the Fisher exact test. This was therefore used to compare the wildtypes and heterozygotes for ascorbic acid and protein in males, females and in both sexes combined.

Although the Fisher's Exact test is considered by some to be overly conservative, there is general consensus in the research community that this test is preferred to the Chi-square test for small data sets. (http://www.childrens-mercy.org/stats/). The numbers of individuals tested were low for the comparison of frequency data however, even with the Fisher's exact test – so these data should be treated with caution. The p values for these tests were therefore calculated with the Fisher Exact test, using R version 2.7.0 (The R Foundation for Statistical Computing). R Code is shown in appendix 5.2.

Ascorbic Acid

Difference between heterozygote Foxf2W174R mutants and wildtypes for ascorbic acid concentrations in urine:

Males only = No significant difference (p-value = 0.0606)

Females only = No significant difference (p-value = 0.2424)

Males & Females = No significant difference (p-value = 1)

Protein

Difference between heterozygote Foxf2W174R mutants and wildtypes for protein concentrations in urine:

Males only = **Significant difference** (p-value = 0.04928)

Females only = No significant difference (p-value = 0.1818)

Males & Females = No significant difference (p-value = 0.06437)

5.8 Intra-Peritoneal Glucose Tolerance Tests

The glucose tolerance test is a means of determining the efficiency of glucose clearing from the blood. The animals were fasted for 16 hours before the administration of an intra-peritoneal glucose injection. Blood glucose was measured immediately, then at 60 minutes and 120 minutes after the injection.

No significant difference between wildtypes and heterozygotes was seen in glucose clearance at any of the time points from the initial tests. P values from T-test comparisons of wildtypes and heterozygotes ranged from 0.4 to 1. A Two way ANOVA with repeated measures was also performed – although there was a data-point missing for the male cohort and it is not generally recommended to perform this type of analysis without a balanced dataset. This gave a p value of 0.97 for males and 0.3 for females, so the lack of correlation between mutant and wildtype IPGTT results was clear for both tests.

The values for each individual that underwent this test were unusually low compared to the expected range at each stage of the experiment. This suggests there is a possibility that insufficient amounts of blood were extracted which would have compromised the data (M.Goldsworthy, personal communication). The IPGT tests have since been repeated by M.Goldsworthy & L. Moir of the MRC Harwell diabetes group with a larger cohort of mice. These additional experiments exhibited a trend towards protection from insulin resistance when fed a high fat diet, when

measurements were taken at 12 weeks, although these results were not significant (as shown in figure 5.10). However tests at 16 weeks on the same mice showed no discernable difference.

One possible explanation for the apparent failure of the original tests to detect a difference between the cohorts is the low number of test animals used for this experiment.

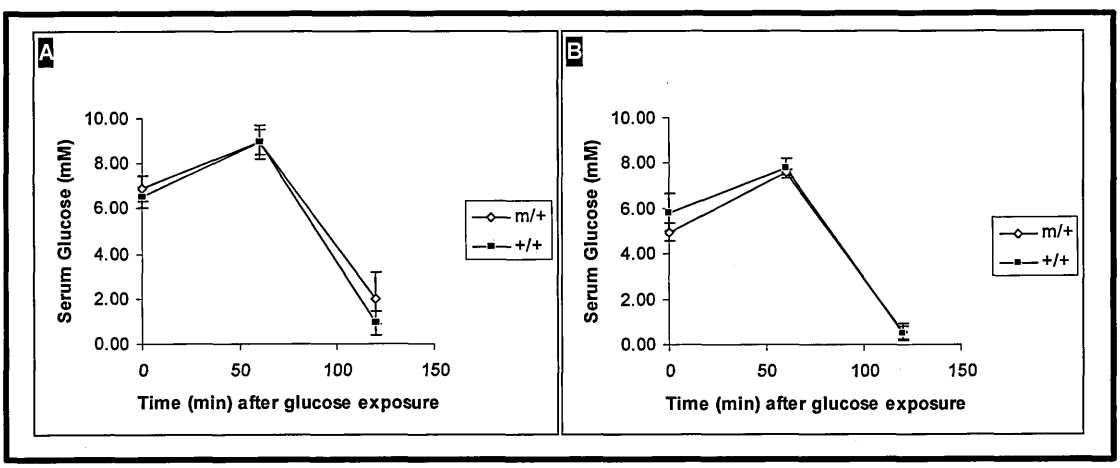


Figure 5.9: IPGTT Preliminary testing
Response to Intra Peritoneal Glucose Tolerance Test by male (A) and Female (B) Foxf2 Foxf2W174R heterozygotes (m/+) and their wildtype littermates (+/+).

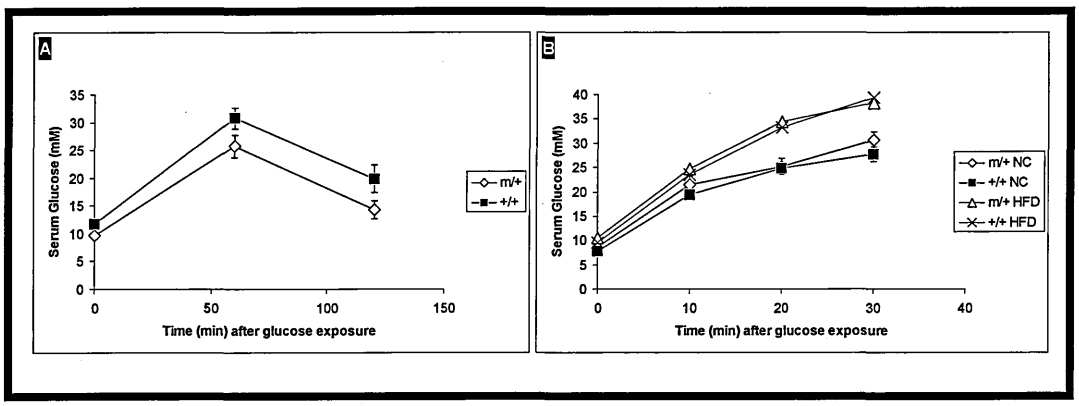


Figure 5.10: IPGTT secondary testing
Reduced Glucose tolerance in Foxf2W174R mice fed on a high fat diet. Intra-peritoneal glucose tolerance tests following a high fat diet in 12 week old (A), and 16 week old (B) wildtype(n=5) and Foxf2^{W174R}(n=7) mice . Foxf2^{W174R} mice demonstrated a trend towards protection from insulin resistance when fed a high fat diet was observed at 12 weeks. Analysis with single factor ANOVA showed p values of 0.065, 0.106 & 0.083 at successive time points in the 12 week old pups. This work was carried out by Michelle Goldsworthy and Lee Moir of the Diabetes group at MRC Harwell.

5.9 Weight Screen

The failure of the Foxf2^{W174R} homozygotes to thrive or to gain weight meant that analysis of the weight gain of heterozygotes during their early post-natal development was of interest.

11 female heterozygotes, 16 female wildtypes, 16 male heterozygotes and 12 male wildtypes were weighed every 3 days from birth until they reached 21 days. Following this, they were weighed every 7 days until they reached 49 days of age (see figure 5.11). Weight measurements until day 21 were obtained by the animal technicians of building 371. ANOVA with repeated measures was performed on the data. This demonstrated that there was no significant difference between the heterozygotes and wildtypes in either the male or female groups.

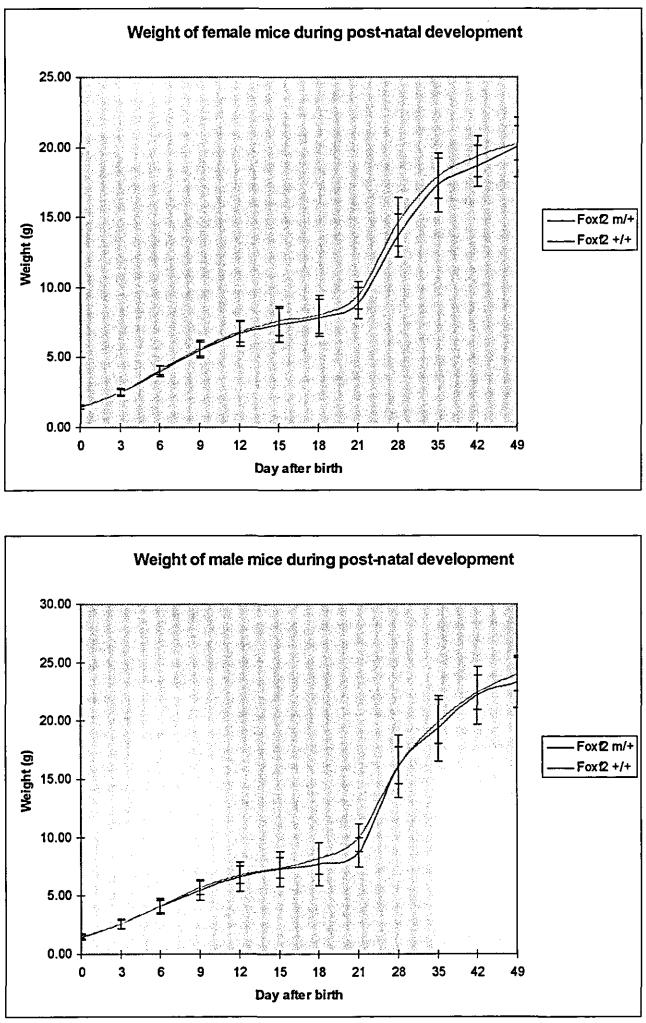


Figure 5.11: Heterozygote weights
Respective weights of heterozygote and wildtype pups between birth and P49. ANOVA with repeated measures was performed on this data: No significant difference between heterozygotes and wildtypes was observed in either female (top graph) or male (bottom graph) weights up to 49 days from birth.

5.10 Eye Analysis

Development of the Ocular Anterior Segment

The anterior segment of the eye is critical for correct ocular function not only because it is the region through which light must pass before contacting the lens, but also because many of its structures are essential for the correct regulation of hydrostatic pressure within the eye. Spaces between the cells of the sponge-like trabecular meshwork, allow the movement of fluid into the canal of Schlemm, from which drainage of aqueous humour into the surrounding venous system can occur [193].

The canal of Schlemm is located at the iridocorneal angle – the site at which the iris, cornea and ciliary body meet. Therefore the correct morphology of this region is critical for regulated drainage – and this is why aberrant development of the iridocorneal angle can cause an increase in intra-ocular pressure.

The ciliary body contains muscles which control the aperture of the iris and therefore the amount of light allowed to pass, via the lens - to the retina. The contraction and expansion of the ciliary muscles can also affect the drainage rate from the adjacent canal of Schlemm. Ciliary processes, the inward folding layers of the ciliary body – also secrete aqueous humour into the anterior chamber, so this feature is key to the regulation of intraocular pressure.

As can be seen in figure 5.12, many of the tissues that are responsible for the control of intra-ocular pressure including the ciliary body, trabecular meshwork and canal of Schlemm are all formed from the periocular mesenchyme. Inappropriate positioning of either the iris stroma or corneal endothelium, which are also of periocular mesenchymal origin, can also impede drainage. It is these tissues which are affected in iridogoniodysgenesis and the anterior segment anomalies.

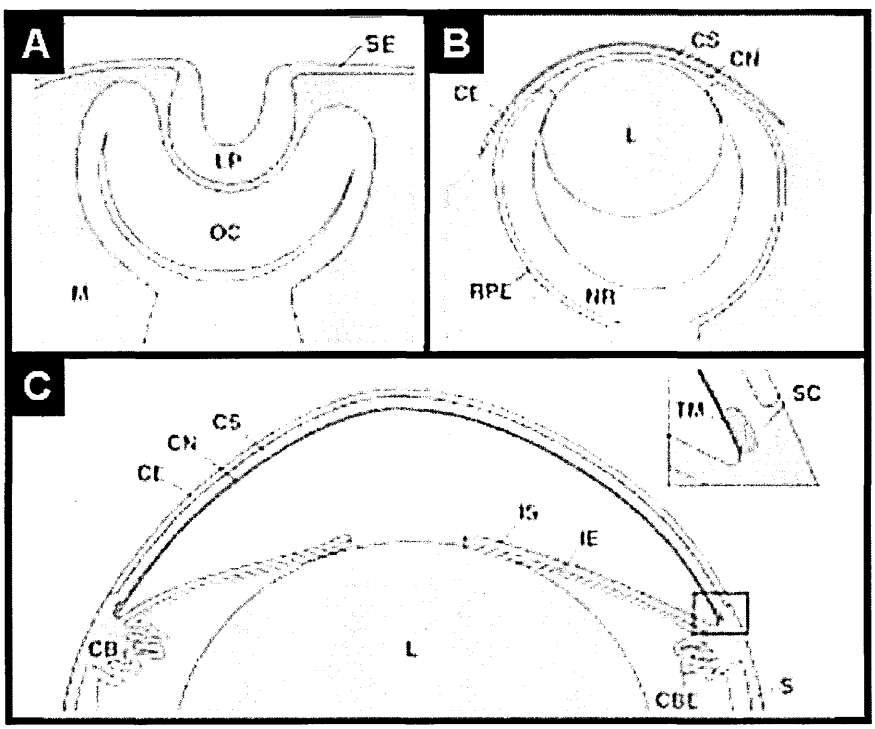


Figure 5.12: Eye development
Surface ectoderm (SE) at the lens placode (LP) invaginates into the surrounding mesenchyme (M) to form a vesicle that separates from the eye-field ectoderm which is destined for the corneal epithelium (CE), to become the lens (L). This vesicle is formed following contact with the optic cup (OC) which is of neuroectodermal origin (NE). The optic cup folds around the lens to form the neural retina (NR) and retinal pigmented epithelium (RPE). These tissues are precursors for the epithelia of the iris (IE) and ciliary body (CBE) in the anterior segment of the adult eye, in addition to forming the retina which is posterior to the lens. The periocular mesenchyme meanwhile, forms the iris stroma (IS), corneal stroma (CS), ciliary body (CB), trabecular meshwork (TM), and the canal of Schlemm (SC). Diagram adapted from Gould et al [193]

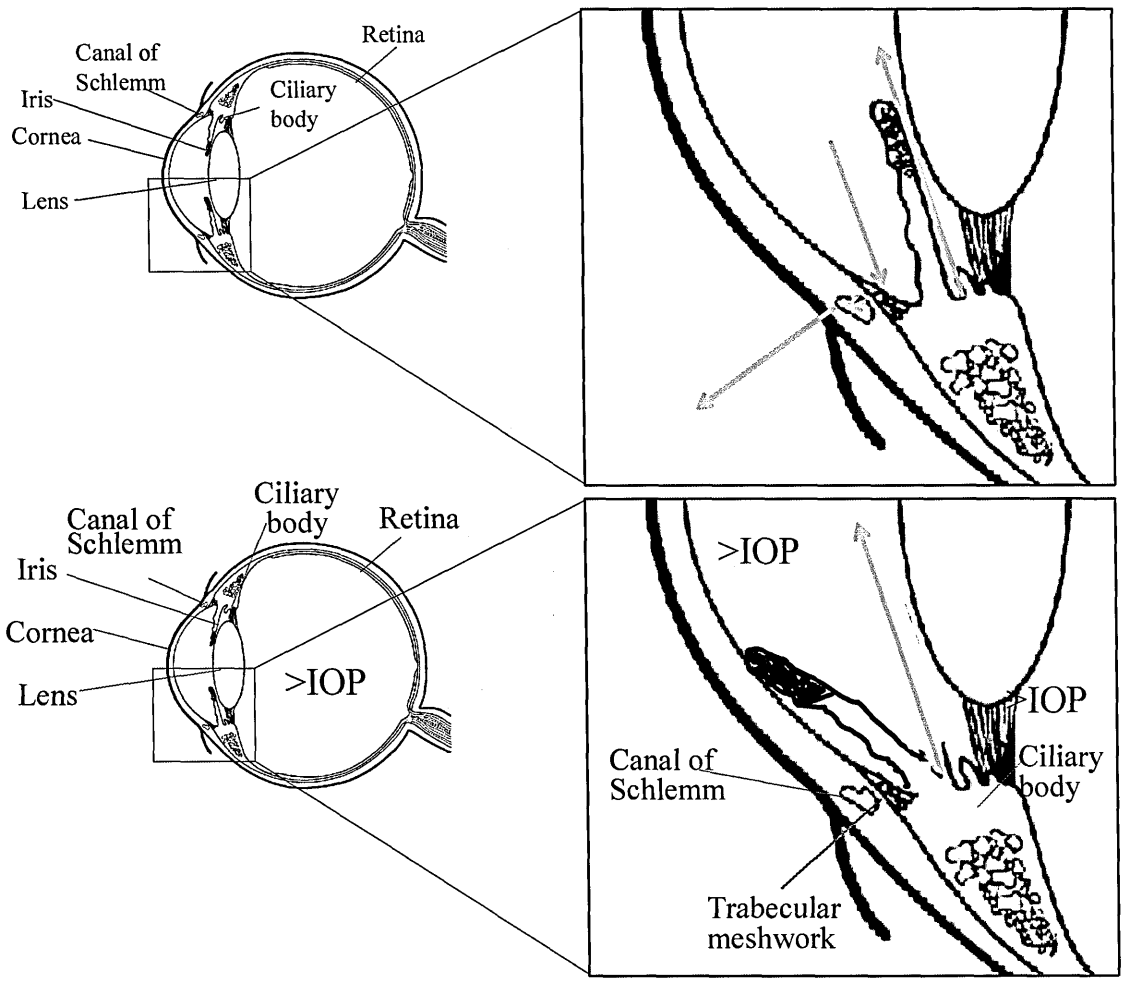


Figure 5.13: Anterior segment dysgenesis can cause an increase in intraocular pressure
Cartoon indicating how a closure of the angle between the iris and the cornea could potentially disrupt the flow of aqueous humour into the trabecular meshwork and ocular drainage system. This would increase intraocular pressure (IOP) and could lead to damage to the optic nerve and retinal nerve fibres. Diagram adapted from medocs.ucdavis.edu/ cha/402/pix/others/eye.gif.

Preliminary analysis of eye histology from Foxf2^{W174R} heterozygotes began with the enucleation and fixation of eyes from 10 heterozygotes (6 females and 4 males) and 4 wildtypes (2 females and 2 males) as described in section 2.2.27. Each of these samples underwent histological analysis by C. Gregory Evans. This analysis was performed with prior knowledge of which tissue came from mutation carriers.

The mice used for this analysis were progeny from a G4 Foxf2^{W174R} heterozygote female and a C57BL/6 male. This cross was required because the inbred mouse strain C3H, which was the predominant background of Foxf2^{W174R} heterozygotes at G4, carries the recessive retinal degeneration allele, Pde6b^{rd1} [262, 263]. Homozygotes for rd1 exhibit a gradual deterioration of vision which could confound any phenotyping of the eyes or visual acuity of the Foxf2^{W174R} mutant. Retinal degeneration would affect analysis of a potential glaucoma locus as retinal damage is one of the symptoms caused by an elevation of intraocular pressure in glaucomatous eyes. Foxf2^{W174R} mutants on a background containing C57BL/6 mice have henceforth been referred to both by the number of generations since the founder male (G1-8), and the number of outcrosses since the C3H line (N1-4).

Following genotyping for Foxf2^{W174R}, each individual was therefore genotyped for rd1 as described in section 2.2.5 and as expected, each was heterozygous for that allele. The recessive nature of rd1 meant that the presence of this allele in a heterozygous state would not confuse phenotypic results.

General morphological features of the regions of interest were collectively analysed in their genotype subgroups (i.e.: regions where a phenotype was observed in the Foxf2 knockout and Foxc1 mutants).

The iris stroma of several mice showed an irregular thinning of the tissue (see figure 5.14), which is characteristic of a transillumination defect in humans, a common trait

of irido-goniodysgenesis. This condition is detected when a narrow beam of light is focused though the pupil at an acute angle, thereby illuminating the retina. Regions of iris thinning will become illuminated by the reflected light when present.

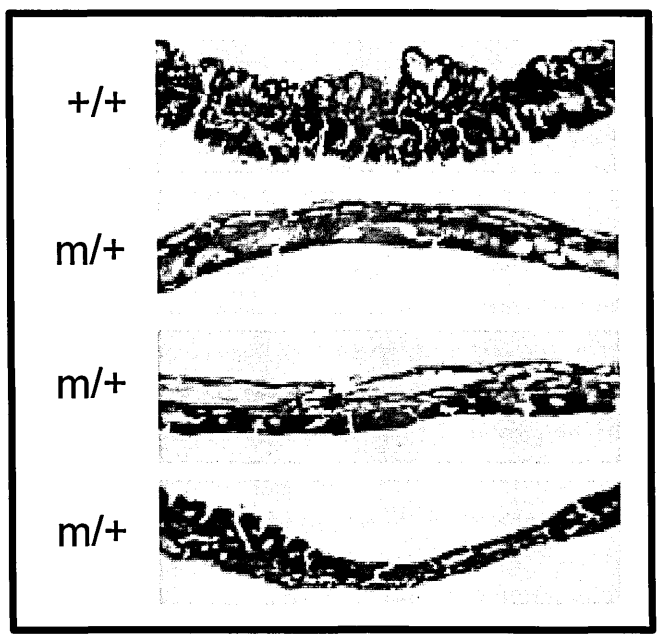


Figure 5.14: Iris stroma malformation
A comparison of wildtype (+/+) and heterozygous (m/+) Foxf2^{W174R} mutant iris stroma. Each of the heterozygote stromas (the three lower images) possess a flattening of the tissue and an apparent reduction of structural integrity. The lowest image shows marked variation in tissue thickness and a loss of structural organisation (in common with the transillumination defect that is seen in human anterior segment anomalies)..

A number of unusual features were observed in the irido-corneal angle of several individuals. The canal of Schlemm was smaller in some of the individuals and was not seen at all in others; the trabecular meshwork showed signs of hypoplasticity; One individual had a hypoplastic ciliary muscle (see figure 5.15).

In several individuals the angle between the cornea and iris was significantly reduced and in one individual the two tissues appear to have fused together. This trait, particularly when associated with the defects in other tissues involved in the

ocular drainage process, could inhibit the outflow of aqueous humour and result in an increase in intraocular pressure. The trabecular meshwork is the point at which the aqueous humour first enters the drainage system and this is located at the irido-corneal angle, which is at the base of both the iris and cornea. A closure of this angle could result in reduction of fluid flow as illustrated in figure 5.13.

Malformation of either the trabecular meshwork or the canal of Schlemm, could also disrupt the drainage of aqueous humour and increase intra–ocular pressure.

A variable phenotype was observed between mice - not all of the eyes had a thinned iris and not all eyes had Schlemm canal or trabecular meshwork defects. Some of the mutants that were tested exhibited both defects, but some only exhibited one.

Key

White arrows = ends of Schlemm's canal

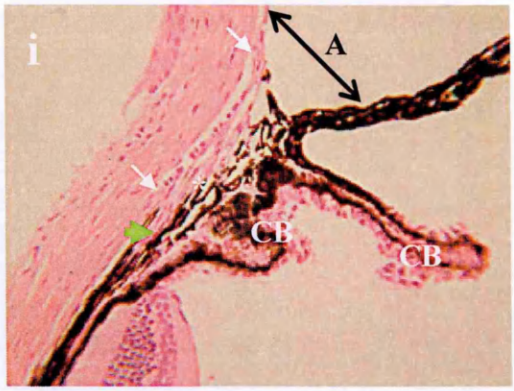
White * = Trabecular Meshwork

CB = Ciliary body processes

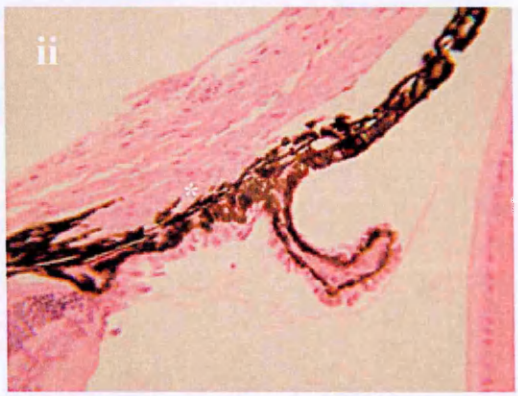
A = Normal irido-corneal angle

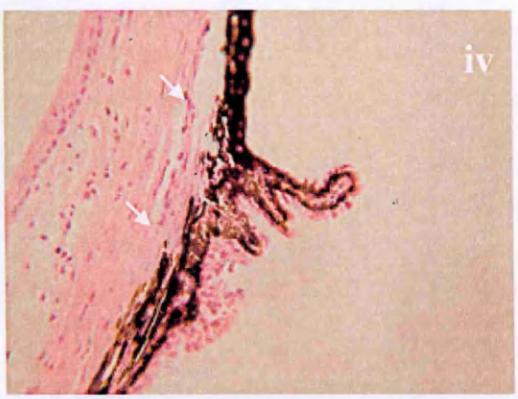
Green arrowhead = Ciliary muscle

i = wildtype



ii - v = Foxf2(174R) heterozygotes





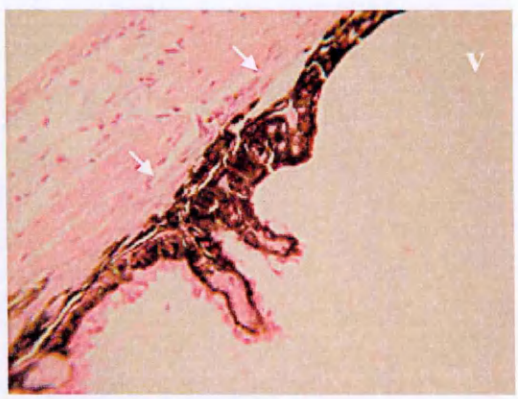


Figure 5.15: Iridocorneal Angle Malformation
(i) is a section from the iridocorneal angle from a wildtype mouse. ii-v are from Foxf2^{W174R} heterozygotes. The Schlemm's canal (arrows) are short or absent in all images. The trabecular meshwork is absent or hypoplastic in all images. The iridocorneal angle is not properly formed, and the iris is attached to the cornea in some cases (v). The ciliary muscle is hypoplastic. in image iii. There is no trabecular meshwork (TM), but there are cells with large nuclei resembling mesenchyme from which the TM is derived - no Schlemm's canal (SC) in this image. (Images supplied by C. Gregory Evans)

The structural changes which were seen in the mutant suggested that a rise in intraocular pressure could occur in these individuals, with the potential to cause damage of the cornea, optic nerve and retinal nerve fibres. Histological analysis showed no sign of any such symptoms in the initial cohort of animals. The phenotype which was seen had the potential to cause a cumulative effect however, so further analysis of older individuals was required to test this hypothesis.

IVFs from Foxf2^{W174R} heterozygote males (N1 C57BL/6) and C57BL/6 females were performed to transfer the colony from a standard mouse facility (Building 371), into a mouse facility with individually ventilated cages (Building MLC). Four heterozygote males, one heterozygote female and one wildtype littermate of each sex were examined with an ophthalmoscope at ~6 months after birth, before humane culling and histological analysis. The histology involved examining sections for retinal and optic nerve defects in addition to looking for the phenotypes seen in the previous cohort. As in the previous cohorts, the wildtypes were phenotypically normal and the heterozygotes showed a range of iridogoniodysgenesis effects. Two male and nine female B6 N1s from the MLC were also examined by ophthalmoscope and histologically examined at 310-350 days after birth; and another cohort of 4 male and 3 female B6 N2s at 400 days old. Of these two aged cohorts, three of 18 mutants demonstrated a thinning iris stroma. Two of the female N1s showed possible eye bulging, which is a symptom of raised intraocular pressure. One of these eyes revealed an unidentified amorphous material between the retina and the lens (figure 5.17).

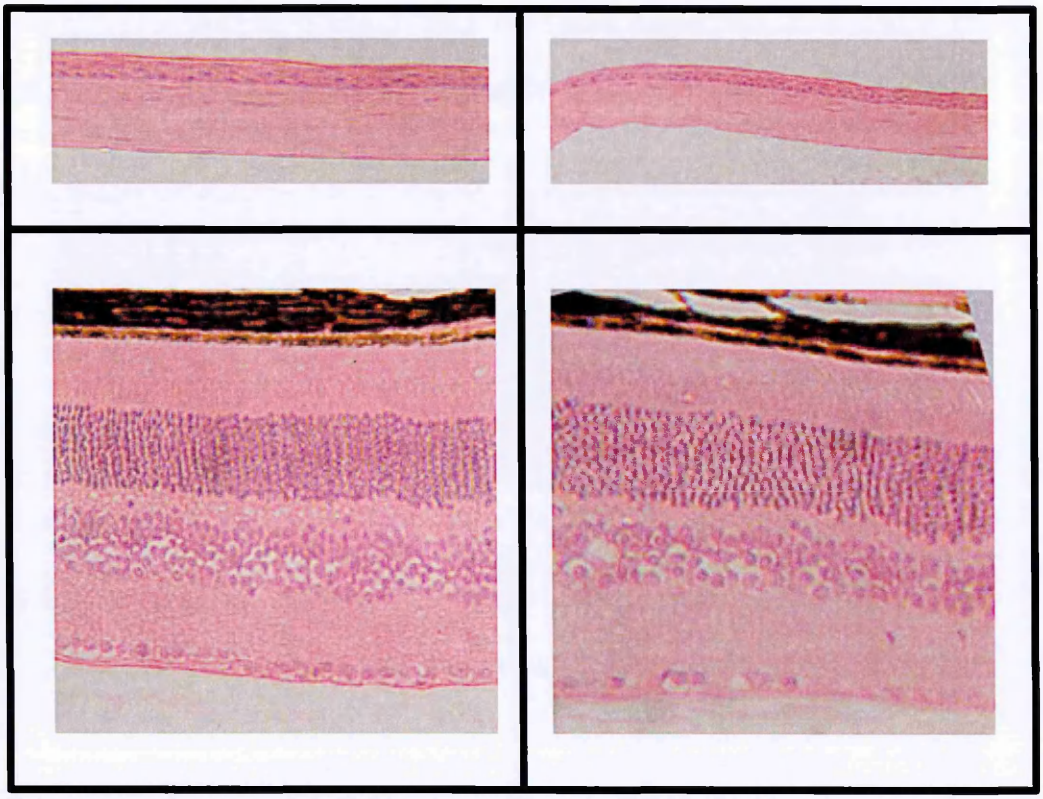


Figure 5.16: Retinal and Corneal Sections from Foxf2W174R Heterozygotes.
As in the younger individuals, no gross abnormalities were seen in either the cornea (top) or retina (bottom) of aged $Foxf2^{W174R}$ heterozygote mice

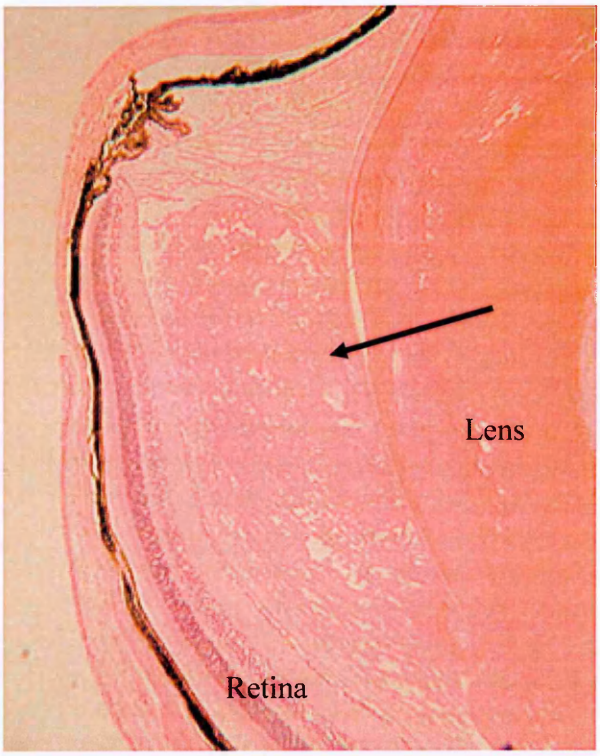


Figure 5.17: Bulging Eyes
Two 318 day old females appeared to have bulging eyes on analysis with an ophthalmoscope, a characteristic trait of raised intraocular pressure. One of the individuals had eyes that contained an unidentified amorphous material between the lens and the retina (arrow) in the posterior chamber, which is occupied by clear aqueous humour in phenotypically normal individuals. This phenomenon was not seen in any other individuals.

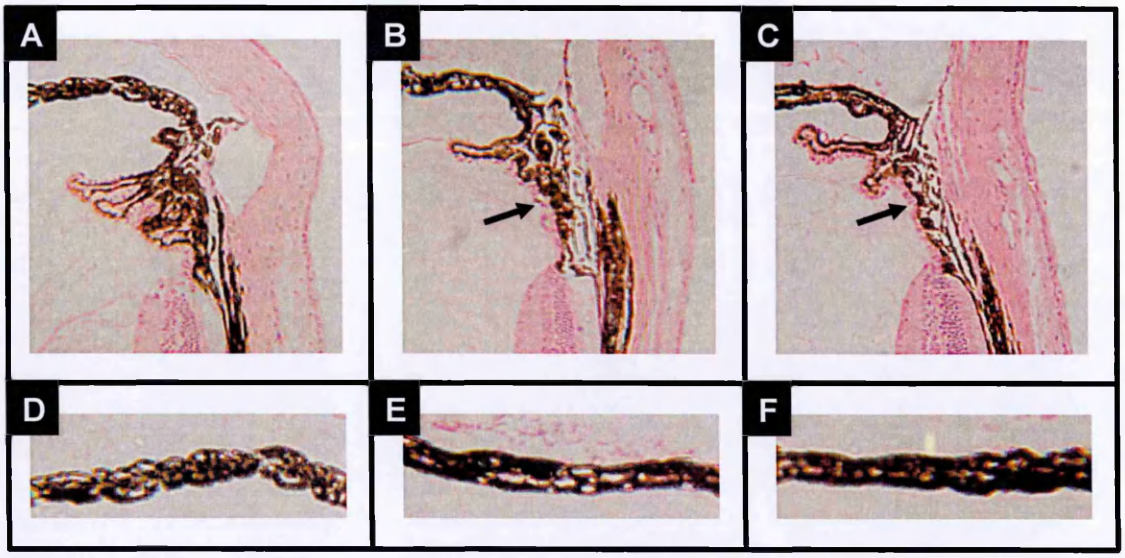


Figure 5.18: Ciliary Body
Some Foxf2^{W174R} aged individuals (B&C) exhibited underdeveloped ciliary bodies, lacking projections towards the retina (arrows) that are present in wildtype eyes (A). A thinned, flattened iris stroma (E&F) as seen in the younger individuals and lacking the structural complexity of the wildtype (D). Three of the 18 heterozygote mice that were analysed at this stage demonstrated thinning of the iris stroma. There were no additional cumulative effects in the older individuals, suggesting that if the pressure within the eye is elevated, the increase is not sufficiently high to cause glaucoma symptoms.

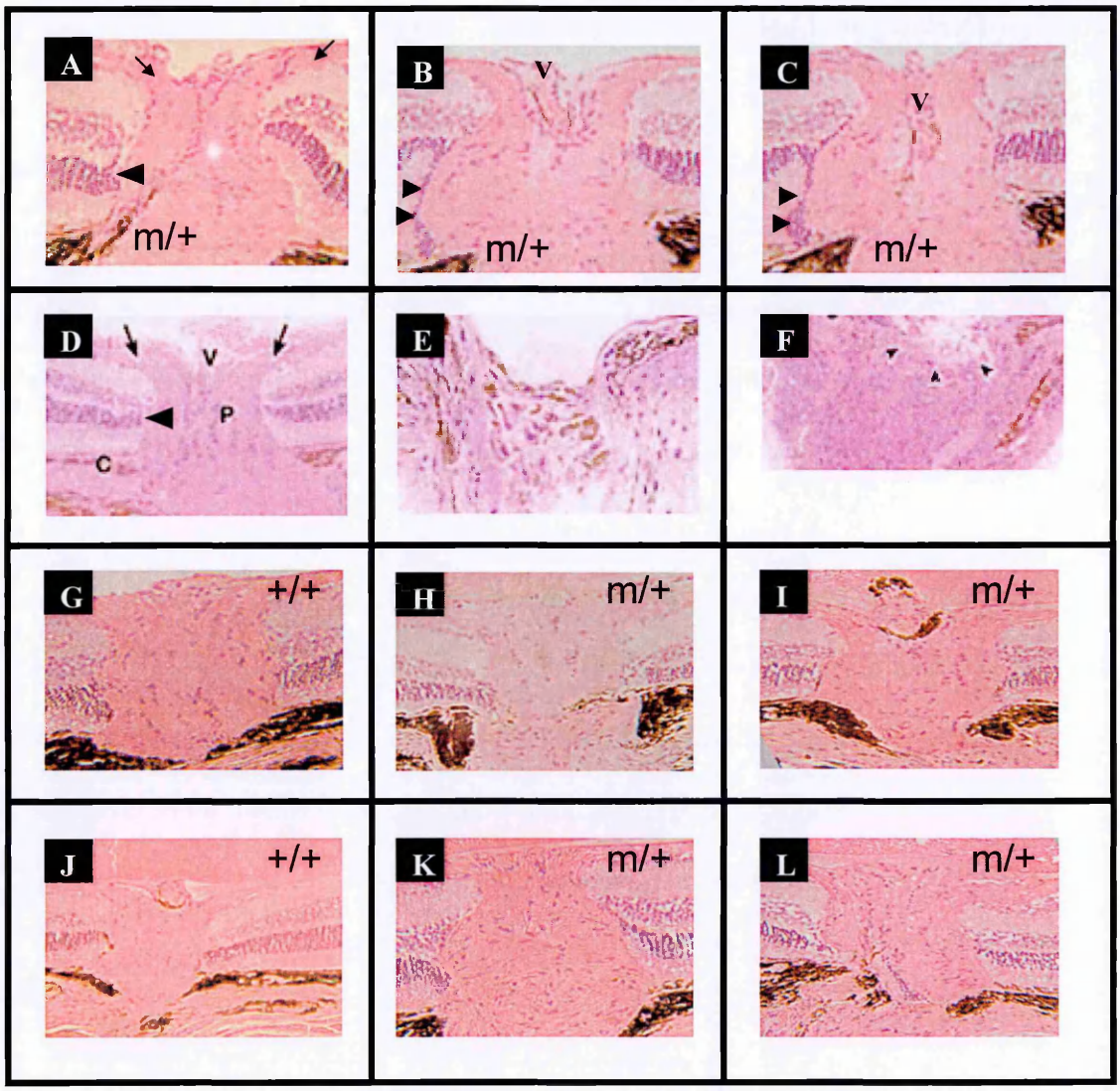


Figure 5.19: Foxf2^{W174R} heterozygote Optic Nerves
Optic nerve sections from Foxf2^{W174R} heterozygote (m/+), wildtype (+/+) and glaucomatous individuals at ~180-540 days. Heterozygotes (A, B & C) at 6 months, glaucomatous eyes (E &F), and a previously published wildtype eye (D). Image A Shows a Foxf2^{W174R} eye which contains neural tissue entering the optic nerve region as occurs in phenotypically normal individuals (D). Images B and C show two sections through the optic nerve of a different Foxf2^{W174R} individual that exhibits a bulge in the nerve fibre layer. (arrow heads). The retina can be seen to continue at the surface of this bulging tissue, whereas normal retinal tissue terminates at the junction with the optic nerve (A and D). V=Vein. C=Choroid. P=Pias. Image E is from an 18 month old mouse with advanced glaucoma, demonstrating excavation of the optic nerve head, a phenomenon which was not present in any of the Foxf2^{W174R} individuals. Image F highlights "cupping" at the optic nerve head in a 12 month old individual. Again, this was not seen in any of the Foxf2W174R individuals. The optic cups of a wildtype 399 day old B6 N2 female (G) and 344 day old B6 N1 male (J) is shown in comparison with a 399 day old B6 N2 heterozygote individual (H) and 310-320 day old B6 N1 female heterozygotes (K, I & L). None of the Foxf2W174R heterozygotes from either of the older N1 or N2 cohorts demonstrated any signs of damage to the optic nerve head.

Eye Screen of heterozygotes aged 330-340(days)

Histological examination was performed on heterozygotes at 330-340 days by Michael Cheeseman during a general screen of their pathological condition. There was no evidence of glaucoma induced injury to the eye. The sections were not perfectly transverse through the optic nerve, but there was definitely no gross enlargement, exopthalmos, injury to the cornea, or disruption of the retinal layers suggestive of prolonged intraocular pressure.

5.10.2 Eye Histology of Homozygote Embryos

E18.5 embryos showed a clear abnormality in the forming ciliary body (See figure 5.20) that was identified during analysis by C. Gregory Evans. There is a complete absence of evagination of tissue in the homozygotes – which can be clearly seen in their wildtype littermates. This evidence supports the hypothesis that the Foxf2^{W174R} mutation influences ciliary development. The parents of these matings were screened to ensure that no rd1 alleles were present, in order to avoid rd1 homozygosity in the Foxf2^{W174R} mutant offspring.

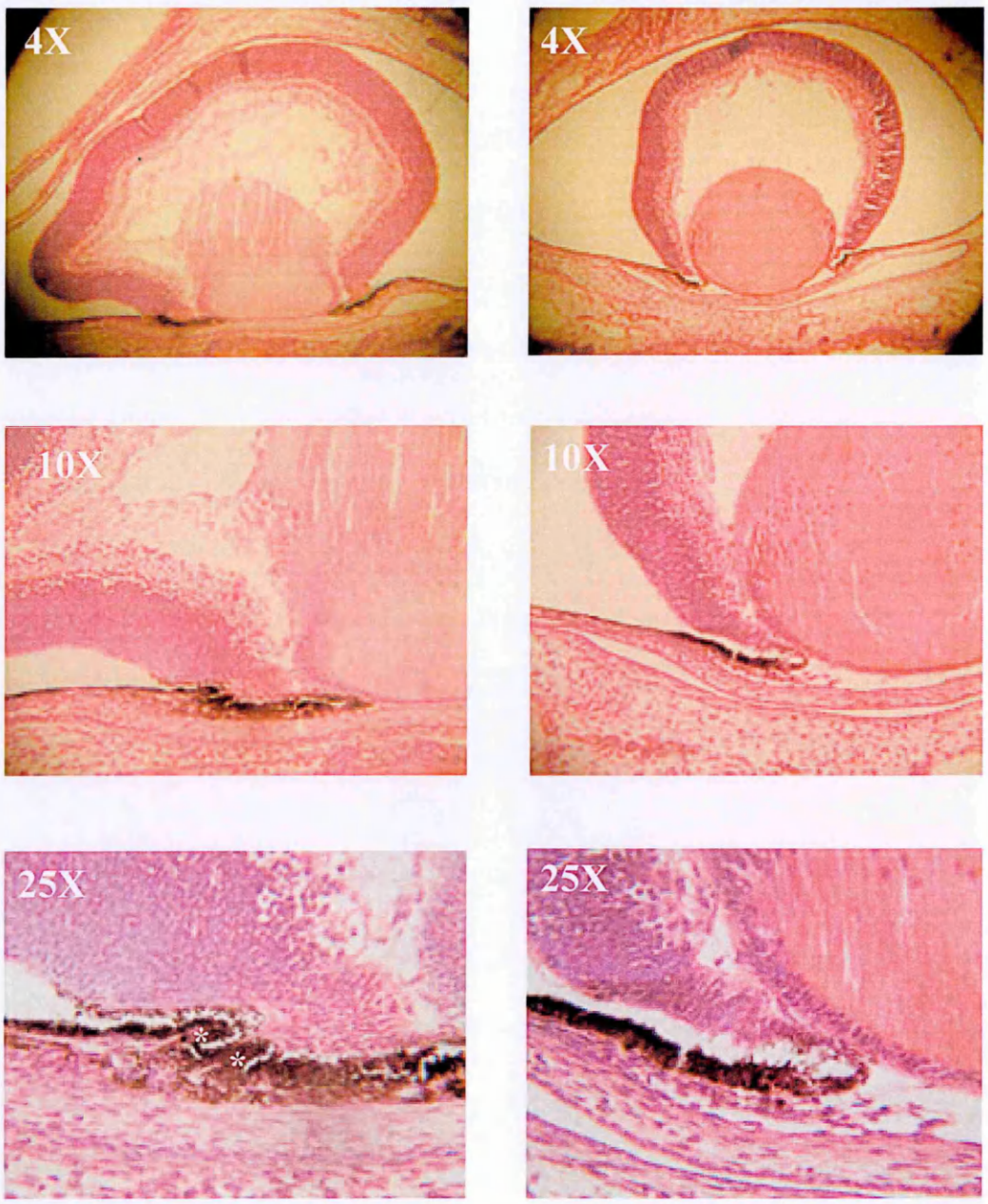


Figure 5.20: Formative ciliary body
The ciliary body is forming in the E18.5 wildtype (+/+) pigment epithelium which has two fingers of evaginating tissue. This was not observed in Foxf2^{W174R} homozygote (m/+) embryos at the same developmental stage. These images suggest that the initiation of ciliary body development is delayed in Foxf2^{W174R} eyes.

5.11 Foxf2^{W174R} Homozygote Analysis

In a preliminary study to examine the effect of the homozygous mutation, a number of G4 intercross matings were set up. Of the initial 3 homozygous individuals which were born, one survived until 13 days and one died at 3 days old. The remaining homozygote was culled at 3 days for pathological analysis. Homozygotes appeared normal at birth, but their condition deteriorated over several days – they did not appear to gain weight at the same rate as their wildtype and heterozygote littermates, and developed dry scaly skin. Hair failed to appear in the individual which died at 13 days although it was present in littermates by 4 days. Milk was seen in very small quantities in the stomachs of some individuals, but was not observed in others. It has been suggested that the homozygote individual shown in figure 5.21C may have developed Chylous Ascites (S. Ball, personal communication), which is the leakage of lymph fluid into the peritoneal cavity.

5.11.1 Genotype Ratios

The wildtype(+/+):heterozygote (m/+):homozygote mutant (m/m) ratio was 16:28:8 in the first Foxf2^{W174R} intercrosses. When this ratio was tested with a chi-squared test; $X^2 = 2.77$ and $p=0.25$ at 2 d.f. (degrees of freedom). This suggests that there is a ~25% probability that chance could cause the observed genotype ratio. Therefore lethality of mutant embryos is unlikely to have contributed to the lower numbers of observed homozygotes.

The number of still-births observed from intercrosses was 1:2:1 which, although too low for any meaningful statistical analysis, is the exact ratio expected if the genotypes varied according to chance alone.

Genotype data from openings carried out at different stages was used to further investigate the possibility of lethality during development in the Foxf2^{W174R} line (table 5.8). The number of live pups at 18.5 days, the stage at which the majority of openings were carried out, was 44 wildtypes: 58 heterozygotes: 34 homozygotes. Again, the chi-square test did not suggest a deviation from the expected ratio of 34:68:34 with a X^2 of 4.41 at 2d.f. which produces a p-value of 0.11.

X^2 and the associated p values were calculated with Microsoft excel.

The Exact binomial test was used to determine whether the heterozygote : wildtype ratios were as expected following Foxf2^{W174R} outcrosses in the Mary Lyon Centre (table 5.10). The calculation was performed using the online binomial test calculator at http://faculty.vassar.edu/lowry/binomialX.html. None of these ratio's were significantly different from the expected ratio of 1:1, suggesting that the Foxf2W174R mutation does not result in an embryonically lethal phenotype – even at low penetrance.

Mouse Cohort	Parental Background					
	All		B6		C3H	
	Frequency	p=	Frequency	p=	Frequency	p=
Total	136		102		34	
Foxf2 +/+	44	0.11	32	0.23	10	0.20
Foxf2 m/+	58		46		12	
Foxf2 m/m	34		21		12	
At 18.5 days:	91		79		12	
Foxf2 +/+	32	0.07	28	0.07	3	0.78
Foxf2 m/+	39		34		5	
Foxf2 m/m	19		15		4	
At16.5 days:	23		23			
Foxf2 +/+	4	0.76	4	0.76	NA	NA
Foxf2 m/+	12		12		NA	
Foxf2 m/m	6		6		NA	
14.5, 15.5, 17.5 days	22				22	
Foxf2 +/+	7	0.22	NA	NA	7	0.22
Foxf2 m/+	7		NA		7	
Foxf2 m/m	8		NA		8	

Table 5.8: Live embryo frequencies
Genotype Frequencies of each live embryo, primarily from openings at E16.5 and E18.5. Foxf2W174R heterozygotes are represented by "m/+"; wildtypes are represented by "+/+". Data from C3H mice at E14.5, E15.5 and E17.5 has also been added, although they are not shown individually due to the low number of openings carried out at these stages – and can only therefore be used to estimate mutant mortality prior to E 14.5. No significant difference between observed numbers and the expected ratio born to each genotype on comparison by chi-squared test.

Sex	Genotype	Background					
		B6N1	B6N2	B6N3	B6N4	C3H G7	C3H G8
Female	+/+	14	12	12	10	13	11
	m/+	8	7	12	9	13	8
	p =	0.29	0.36	1.00	1.00	1.00	0.65
Male	+/+	15	17	15	11	14	13
	m/+	10	14	12	9	15	15
	p =	0.4	0.7	0.7	0.8	1.0	0.9

Table 5.9: Frequencies of all live births in the Mary Lyon Centre.
Foxf2W174R heterozygotes are represented by "m/+"; wildtypes are represented by "+/+". C3H mice are at generations G7 and G8 with respect to the founder male. B6 N2, N3 and N4 mice were derived from a preliminary outcross of C3H mice at G4 (and are therefore G6, G7 and G8 generations of outcrosses from the founder male). The B6N1 mice were the offspring of C3H carriers at G6, so were at G7 with respect to the founder male.

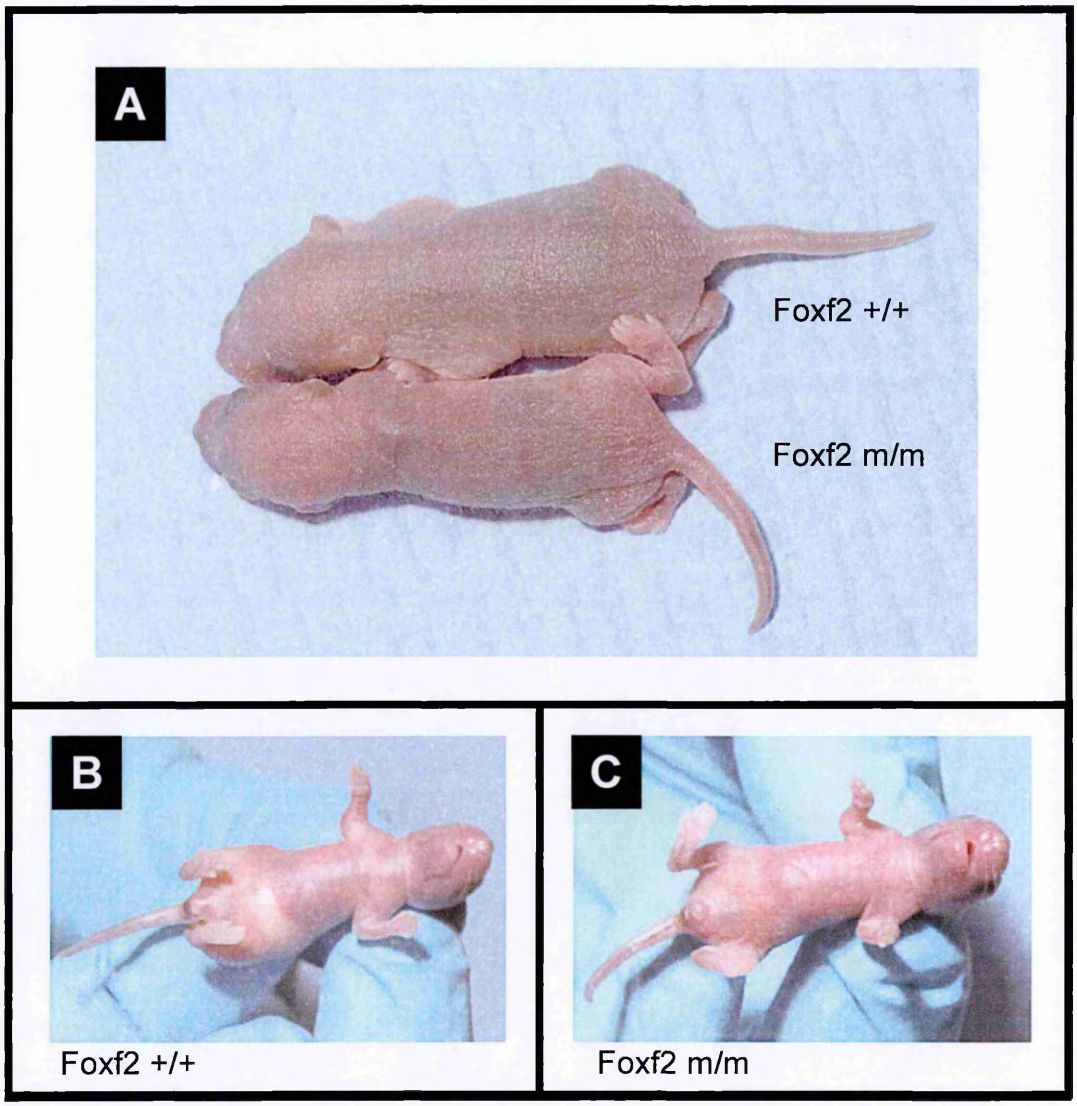


Figure 5.21: Homozygous pup
A three day old wildtype (+/+: upper pup on A and pup B) and Foxf2 homozygote (m/m: lower pup on A and pup C). There is no evidence of milk in the stomach of the homozygote pictured, although small amounts were observed in some homozygote individuals.

A further set of 5 intercross matings were set up, followed by the live births of 5 confirmed homozygotes from a total combined litter size of 27 live births. One homozygote was born dead. There were also four additional still births which weren't homozygotes. There was therefore a total of 6 homozygotes from 34 births and a ratio of 11:13:5 ((+/+):(+/m):(m/m))in this second batch.

The surviving five homozygotes all showed the same signs of deterioration as seen in the first three homozygote individuals to be born, i.e.: failure to thrive or gain weight, absence or only small quantities of milk in the stomach, lack of hair and at 3 days, dry flaky skin. These were culled before their natural deaths due to project licence restrictions.

The high incidence of still-births born from intercrosses cannot be explained by the Foxf2 mutation alone, because there was only one homozygote amongst the 5 still-births which occurred. It seems likely that the higher incidence of stillbirths in intercross matings was the result of additional ENU induced mutations. If this is the case, intercrosses from later generations would be expected to result in lower numbers of still births because a smaller percentage of the parents' genome would originate from the mutagenised founder male. One alternative hypothesis is that Foxf2W174R heterozygous mothers are not delivering sufficient nutrition to the unborn pups.

Matings to produce $Foxf2^{W174R}$ homozygote births were stopped before the cause of death could be established, due to severity limits imposed by the project licence. The intercrosses were however used to produce homozygous foetuses for analysis of the eyes, lungs and palates.

5.12 Pathology/Necropsy Reports

5.12.1 Heterozygote Pathology Screen

Four Foxf2^{W174R} G7 heterozygote mice (three B6 N1 males and one B6 N3 female) of between 330 and 400 days of age underwent a general pathology screen (carried out by M. Cheeseman) in order to identify potential age-onset conditions associated with the Foxf2^{W174R} mutation. Particular attention was paid to the gut in this screen, following a report that mentioned the thinning of the gut walls in Foxf2 homozygous knockouts [154].

Nothing unusual was found in the guts of these individuals. One individual male had an epithelial cyst of the tongue. The only female to be examined had white specks on her trachea, and on observation under a microscope, this region showed tracheal gland dilation. Another male had mineralisation of the blood vessels of the thalamus. None of these observations were regarded as anything other than age related incidental findings.

5.12.2 Screen of Homozygotes at day 3

The homozygote pup which was culled at 3 days for pathological analysis had developed sinusitis associated with inhaled bedding. It had also developed bronchiolitis with a solitary microabscess of the lung, accompanied by inhalation of milk. The middle ear had mild otitis media.

Each of these inflammatory changes were classified as mild to minimal, acute onset and associated with a neutrophil leukocyte reaction. The low severity level of each of these conditions meant that they may not be clinically significant and were probably secondary to the primary lesion.

It was suggested in the report that the bronchiolitis and mild otitis media were possibly caused by the accidental inhalation of bedding and milk rather than due to the mutation.

Milk inhalation can be an indicator of a cleft palate, however the low-level milk inhalation in this animal could have been within the normal spectrum expected in animals at three days of age.

There was no dilation of the oesophagus. This was highlighted as evidence against the presence of a cleft palate as those animals with a cleft can have distended stomachs caused by unsuccessful attempts at feeding. A cleft was not identified in this individual. It was suggested by Michael Cheeseman that the inhalation of milk may have been due to a physiological discoordination in the swallowing reflex.

5.13 Cleft Palate Tests

The Foxf2 -/- knockout described by Wang [153], had a severe cleft palate which resulted in a gas distended gut, and died within 18 hours. The cause of death was probably starvation due to an inability to suckle. The comparatively extended lives

of the Foxf2^{W174R} homozygotes, and absence of a distended, air-filled stomach indicated that a severe cleft palate was unlikely in these individuals. However, the reduction in milk uptake and inability to thrive followed by an early death, suggested that the Foxf2^{W174R} individuals could have had a less severe cleft.

Sixteen homozygote individuals were examined for clefts of the primary and secondary palate microscopically with assistance from H. Tatteossian. Confirmation of a cleft palate was not seen in any of these individuals.

Of the homozygote individuals examined, three were from openings at E14.5, one was at stage E15.5 and a further four were at E16.5. These stages cover the period during which the maxillae of the secondary palate are in the process of fusing. Any developmental problems with the fusion process are more likely to be made clear at this stage. No abnormalities in the early stages of fusion were detected, however.

Palates were also examined at the later stages of E17.5 (n=4) and E18.5 (n=4), when the palate should be fully formed. There was no evidence of abnormalities in thoracic contents. None of these homozygotes had a cleft palate in the serial sections investigated, although one individual had tissue damage of the palate that appears to have occurred during sample preparation (figure 5.22). This may indicate a deficiency in the structural integrity of the palate. Anatomically normal incisive foramina, which are a pair of openings into the palate positioned behind the incisor teeth, were present in rostral sections of all individuals.

Figure 5.22: Secondary palate of Foxf2^{W174R} Homozygotes
Secondary palates of three E18.5 individuals: Two Foxf2^{W174R} homozygotes (m/m) and one wildtype (+/+). One homozygote individual was identified with a suspected cleft of the secondary palate (B – the cleft is highlighted by the black arrow). However, analysis of 15 additional homozygote individuals (e.g.: C), failed to identify further individuals with cleft palates. On closer examination, it appears that the cleft seen in image B, may be the result of tearing during histological preparation rather than a developmental malformation, although it is possible that there is a reduction in the structural integrity in this tissue which facilitated this tearing.

5.14 Lung Histology

Whole lungs from 12 embryos and the accessory lobes from 9 embryos were extracted as described in section 2.2.27 with assistance and training from Dr. Charlotte Dean of MRC Harwell. Following the extraction of lung tissue, tail tissue was extracted and digested as described in method 2 of section 2.2.2 for subsequent genotyping and identification of Foxf2^{W174R} homozygotes. The whole lung of one homozygote E18.5 individual and the accessory lobe of the lung from another E18.5 homozygote individual were examined for deformities and compared with wildtype littermates with the assistance of Dr C. Dean. There was nothing unusual in the structure or organisation of the pneumocytes or parenchyma in the heterozygote lungs.

5.15 Prognosis

There remain many questions about the effect of the Foxf2^{W174R} mutation. There seems to be a good chance that the homozygous mutation is responsible for the deaths of the pups in the homozygote study but the cause of death remains unclear. Initial suspicions that a subtle cleft palate may be responsible for the deaths have been disproved and there is no obvious defect in the lungs. The focus in any further work to identify the cause of deaths should include a study of the gut, following work carried out by Ormestad which identified severe structural problems and aganglia of the guts of Foxf2 knockout mice[154].

One other possible route would be a detailed analysis of the tongue. In his pathology investigation, Michael Cheeseman suggested that the inhalation of milk (which could have been the cause of both the bronchiolitis and the otitis media) may have been due to a "physiological discoordination in the swallowing reflex". Normal function of the tongue is necessary for controlled swallowing, and the malformed tongue identified by Wang in the Foxf2 knockout [153] would suggest this as a further route of investigation. However no obvious tongue defects were detected during the screening of transverse head sections for the analysis of palate development.

The irido-goniodysgenesis phenotype in the Foxf2^{W174R} heterozygotes has provided further evidence for the involvement of the gene in glaucoma. No glaucoma was identified in these mice, but because the mutation seems to disrupt the normal

formation of every ocular drainage tissue, it is possible that intraocular pressure is affected. Although not utilised in this study, the measurement of intraocular pressure in mice could be achieved by cannulation of the eyes of an anesthetised individual with a microneedle linked to a pressure transducer as previously described [264].

The initial significant PPI result was not repeatable with either the same group of animals or another smaller cohort. Further tests under alternative volume and frequency conditions with further, larger cohorts did not yield significant results. As such, it does not appear that the preliminary data shows a genuine difference.

The startle effect however, produced highly significant results when the testing was under tightly regulated conditions and when the mice used were less likely to be affected by background heterogeneity. Although not entirely conclusive due to the earlier absence of a significant affect, these results suggest that Foxf2W174R mice appear to demonstrate a genuine elevated startle response.

Differences between heterozygote mutants and wildtypes were detected in calcium ($p<0.05$), potassium ($p<0.01$), and chloride ($p<0.01$) in bleeds taken from 5 month old males and urea ($p<0.05$) in females. However further analysis at 8 months showed no significant difference between the mutants and wildtypes in any of these metabolites. Urea, potassium and chloride can all act as indicators of renal problems. The significant urine dipstick results could also be a sign of kidney malfunction:

ascorbic acid is reabsorbed in the kidney [265] as is serum protein [266] - proteinurea is often associated with renal disorders.

Foxf2 expression was not detected in the mouse kidney by Aitola [188] either during development or in adults, and this absence was specifically mentioned by the author. The antibody company Abcam, recommend the use of kidney tissue as a positive control for their anti-human FOXF2 polyclonal antibody (at http://www.abcam.com/index.html?datasheet=23306). The reliability or perhaps the identity of this antibody was initially called into question, by its failure to detect FOXF2 in Cos-7 cells that were transfected with a FOXF2 construct when other evidence suggested the presence of the protein (see section 6.2.6)). However, this antibody was used successfully by Philip-Couderc to identify the presence of the FOXF2 protein [231]– which was derived from the same plasmid [203] as the one used in the experiments described in chapter 6.

A kidney phenotype is possible even if the gene is not expressed in the kidney. For example, diabetes mellitus is a common cause of kidney disorders due to the damaging effect of high blood-sugar levels on renal blood vessels and may be the result of abnormal expression of genes in other tissues. No difference was seen in the blood glucose levels of mutants but unpublished data from glucose tolerance experiments that were performed by Dr M. Goldsworthy of the diabetes group at MRC Harwell suggest that $Foxf2^{W174R}$ mice have innate protection against insulin resistance when fed a high fat diet.

Chapter 6: Molecular Phenotyping of Two FOXF2 mutations

6.1 Introduction

The Foxf2 protein is a transcription factor and as such, it binds to promoter elements and increases or reduces the expression rates of target genes. Previous reports suggest that it is a transcriptional activator [203, 242]. The mRNA transcript is present in cells of the mesenchyme adjacent to the endoderm-derived epithelium in the gut, genitals and lungs [154] and adjacent to the ectoderm derived epithelium in the tongue [153, 154].

Following the publication of the Foxf2 knockout by Wang et al [153], the characterisation of the physical phenotype was expanded and the molecular effects further characterised by Ormestad and colleagues with the analysis of the knockout and Foxf2/Foxf1 compound heterozygotes. They discovered that Foxf2 played an active role regulating the Bmp, hedgehog and Wnt pathways [153, 154].

Immunohistochemistry showed a drastic decrease in types I and IV collagen and a lesser but still significant reduction in Laminin [154] (see figure 6.1). These proteins are major constituents of the extracellular matrix (ECM) and accordingly there was a severe reduction in the ECM of mutant intestines.

A deficiency of ECM was observed in the intestines of FOXF2 -/- knockouts, and a reduction in collagen formation led to the suggestion that ECM abnormalities may contribute to anterior eye defects in Foxc1 and Foxc2 mutants [81, 138] as well as

mice with mutations in Bmp4 [267] and Tgfb2 [268]. In addition, Foxc1 homozygous knockout mice have ECM abnormalities in other tissues such as the arachnoid layer of the meninges in the central nervous system and the prechondrogenic mesenchyme of cartilage [81], [110]. The analysis of Foxf2 interactions in the gut has highlighted the gene's involvement in extracellular matrix production.

The eye phenotype of the Foxf2^{W174R} mutant has remarkable similarities to the published Foxc1 knockout [81] and it is possible that a deficiency in ECM is the common cause. One or more types of collagen are candidate downstream targets of these genes, and aberrant collagen expression could forseeably cause a deterioration in ECM quality that would result in anterior eye defects like those seen in Foxf2^{W174R} mice.

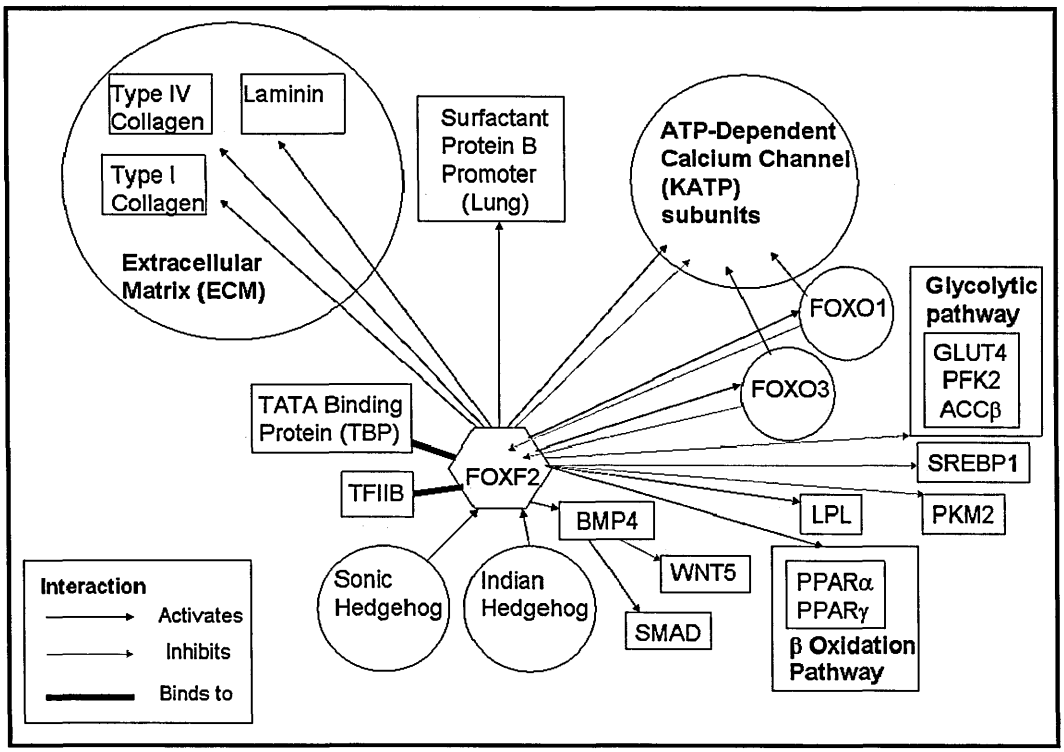


Figure 6.1: Foxf2 Molecular interactions
A general scheme of the known molecular interactions of FOXF2 at the time of writing. Direct and indirect regulatory effects are not differentiated in this figure – many of the developmental interactions were deduced by Ormestad et al [154] from immunohistochemical experiments on Foxf2 knockout mice and Foxf1/Foxf2 knockout compound heterozygotes. It is not yet known whether Wnt5 interacts directly with Foxf2, but the levels of this molecule increase in FOXF2 homozygous knockout mice, suggesting that the forkhead molecule has an inhibitory effect on the Wnt pathway in Foxf2 wildtypes. All the metabolic interactions, including those with the KATP subunits were substantiated by Philip-Couderc et al [231] by immunofluorescence, western blots and quantitative RT-PCR

Gould and co-workers presented a model of ocular anterior segment development, which was inspired by the similar phenotypes of Tgfβ2, Bmp4, Foxc1 and Foxc2 mouse mutants, amongst others. All of these mice contain deficiencies of the ECM in the iridocorneal structures. The authors suggest that cell surface recognition of BMP4 and TGFβ2 triggers a SMAD pathway which is regulated by several suggested molecules including Foxc1 and Foxc2. These transcription factors may control the quantity and structural integrity of extracellular matrix [193]. The same

logic could also be extended to include Foxf2, due to the phenotype in the anterior segment that was observed in the $FOXF2^{W174R}$ mutant (section 5.10).

During intestinal development, *Bmp4* has been identified as a downstream target of *Foxf2* [154]. This contrasts with the suggested Fox/Bmp interaction in the eye model, in which *Foxc1* and *Foxc2* were hypothesised to be upstream regulators of *Bmp4* [193]. It is possible that the Foxf genes occupy a different level of the hierarchy within some of the same pathways as the Foxc genes, but it also remains possible that they perform similar functions in some tissues and are partially redundant.

Following work which demonstrated reduced levels of Bmp4 and increased levels of Wnt5a in the Foxf2 mutant, a possible role can be seen for Bmp4 as an intermediary between Foxf2 and the canonical Wnt pathway [154]. Wnt5a levels were increased in Foxf2 knockout explant cultures but were not increased in these tissues when exposed to Bmp4. An increase of Wnt5a was also seen in wildtype explant cultures that had been exposed to the Bmp4 antagonist noggin, suggesting that Bmp4 inhibits Wnt activity.

6.1.1 FOXF2 Nuclear Localisation & Mode of Action

Like other forkhead-box proteins, FOXF2 is localized primarily to the nucleus of the cell. Hellqvist et al discovered that FOXF2 nuclear localisation depended on sequences from both ends of the forkhead domain [242]. As with other features of the forkhead domain, the highly conserved amino acid sequence means that this may be a widespread trait throughout the gene family and indeed, Foxc1 also contains nuclear localisation signals at the N- and C- termini of the forkhead box [269]. One member of the forkhead family, FOXP2 -appears to be localized to the nucleus but excluded from the nucleoli [270] and again, this may also be true of other forkhead genes but is yet to be investigated in the wider gene family.

In addition to binding DNA and nuclear localization, the forkhead domain also facilitates interaction with the basal transcription machinery by binding to the TATA-binding protein (TBP) and TFIIB [242]. The amino acid sequence at the C-terminal portion of the domain that binds to TBP is highly conserved [205]. This suggests that TBP binding may be another feature that is common to all members of the forkhead box family.

The forkhead domain is responsible for the recognition of target sequences. There is a wide variety of recognition sequences however, and it has been suggested that the amino acids that flank the forkhead domain as well as differences within this domain may be responsible for the different specificities of the forkhead protein family [205]. Fox genes with indistinguishable DNA binding specificities can have distinct biological functions. For example although both FOXF2 and FOXF1 proteins bind

to and activate the surfactant protein B promoter, the gene for another lung-specific protein, CC10, is only activated by FOXF1 due to the presence of a cell type-specific activation domain in this molecule [203].

Two mis-sense mutations in the forkhead domain of Foxc1 described by Saleem et al [271], allowed DNA binding to occur as normal, but disrupted transactivation.

There is strong evidence for an anterior eye defect in the W174R mutant and several additional potential phenotypes that require further investigation. It therefore seemed pertinent to examine the molecular interactions that are responsible for these mutant characteristics.

In addition, although the V412F mutant line has not yet been re-derived at the time of writing, it also seemed worthwhile to replicate the W174R experiments with this mutation as a preliminary investigation into the effect that the mutation has on DNA binding and activational activity. Although not in the binding domain and therefore less likely to effect DNA binding, its position within one of the activation domains could provide useful information about the functional structure of these regions.

The third mutation discovered in chapter 4 occurs within the splice site of FOXF2 immediately adjacent to the 5' end of exon 1. As such, experiments that are designed to analyse binding and activation are not appropriate for this mutation because a more important issue in this case is the splicing efficiency of Foxf2 and total levels

of the intact protein in these animals. An intronic splice site mutation is unlikely to impact on DNA binding

6.1.2 The Foxf2 Binding Sequence

A consensus core binding sequence for four forkhead (previously known as "FREAC" for Forkhead Related Activators) proteins was described by Pierrou et al [205] who performed gelshift assays with random sequence oligonucleotides, and confirmed the consensus sequence by DNA footprinting. This core of RTAAAYA was common to each of these forkhead proteins and there was a complete absence of any bases other than the consensus core in any of the random oligonucleotides tested. However the flanking sequence favoured by each of the proteins varied – as did the purines in the core at position R and pyrimidines at position Y.

The Purine at position R in the core sequence was a guanine in 74% and an adenine in 26% of the oligonucleotides that bound to FOXF2, the position Y pyrimidine was a Cytosine in 93% and a Thymine in 7% of oligonucleotides (see figure 6.25C). The flanking sequence demonstrated a definite preference for particular nucleotides - each of the other bases were observed at these flanking positions but at significantly lower percentages, and Pierrou was therefore able to highlight a binding sequence specific to FOXF2: aacGTAAACAa.

Miura used the FOXF2 protein-DNA binding data obtained by Pierrou et al [205] to design oligonucleotides for a transactivation assay, in which he compared the

activity of different truncated regions of the mouse Foxf2 protein [241] (previously known as "LUN"). He incorporated each of the 16 nucleotides that were examined by Pierrou et al at and near the core sequence, including those bases which Pierrou and colleagues did not consider frequent enough to highlight as significant. This expanded the binding sequence contained within these oligonucleotides to: caaacGTAAACAatcc.

These oligonucleotides have been adapted for a comparison of the activational properties of wildtype and mutant FOXF2 proteins using pTAL-SEAP as a reporter vector (see figure 6.21).

6.2 EMSA

The EMSA (Electrophoretic Mobility Shift Assay) or Band Shift Assay is a method for studying protein-DNA interactions *in-vitro* [272]. Oligonucleotide probes which contain target sequence or candidate target sequence for the protein of interest are labelled with a visualisation molecule before mixing this probe with the protein and running through an electrophoretic gel.

These gels are run under native (non-denaturing) conditions and therefore the charge, morphology and molecular weight of the protein influence its migration. Native gels are necessary because DNA binding is highly sensitive to alterations in molecular morphology and an appropriate interaction depends on the correct alignment of these molecules.

Protein-probe complexes will migrate slower through agarose or acrylamide gels than free probe. Unbound probe is added to some lanes in order to compare its rate of migration with potential probe-protein complexes. Binding of the probe to the protein would result in a reduction in the migration rate through the gel - due to the larger size of the complex than either of the molecules alone.

Antibody specific to a protein of interest is often added to EMSA binding mixtures. If this antibody binds to the protein-DNA complex, then an additional band should be seen with a slower migration rate than either the unbound DNA probe or the DNA-protein complex. This test is known as a "supershift" assay and a positive result acts as unambiguous proof of a proteins identity.

An increasing gradient of unlabelled probe can be added to the binding mixture in several wells as a further control. If the interaction between the probe and the protein is specific, then there should be a sharp reduction in the strength of signal as the concentration of inhibitor increases. Non-specific competitor should require a much higher concentration in order to eliminate the labelled probe-protein complex than the specific competitor.

The oligonucleotides that were initially designed for the pTAL-SEAP/FOXF2 transactivation assay (section 6.3) form a double stranded FOXF2 binding sequence and therefore also suffice as an EMSA probe.

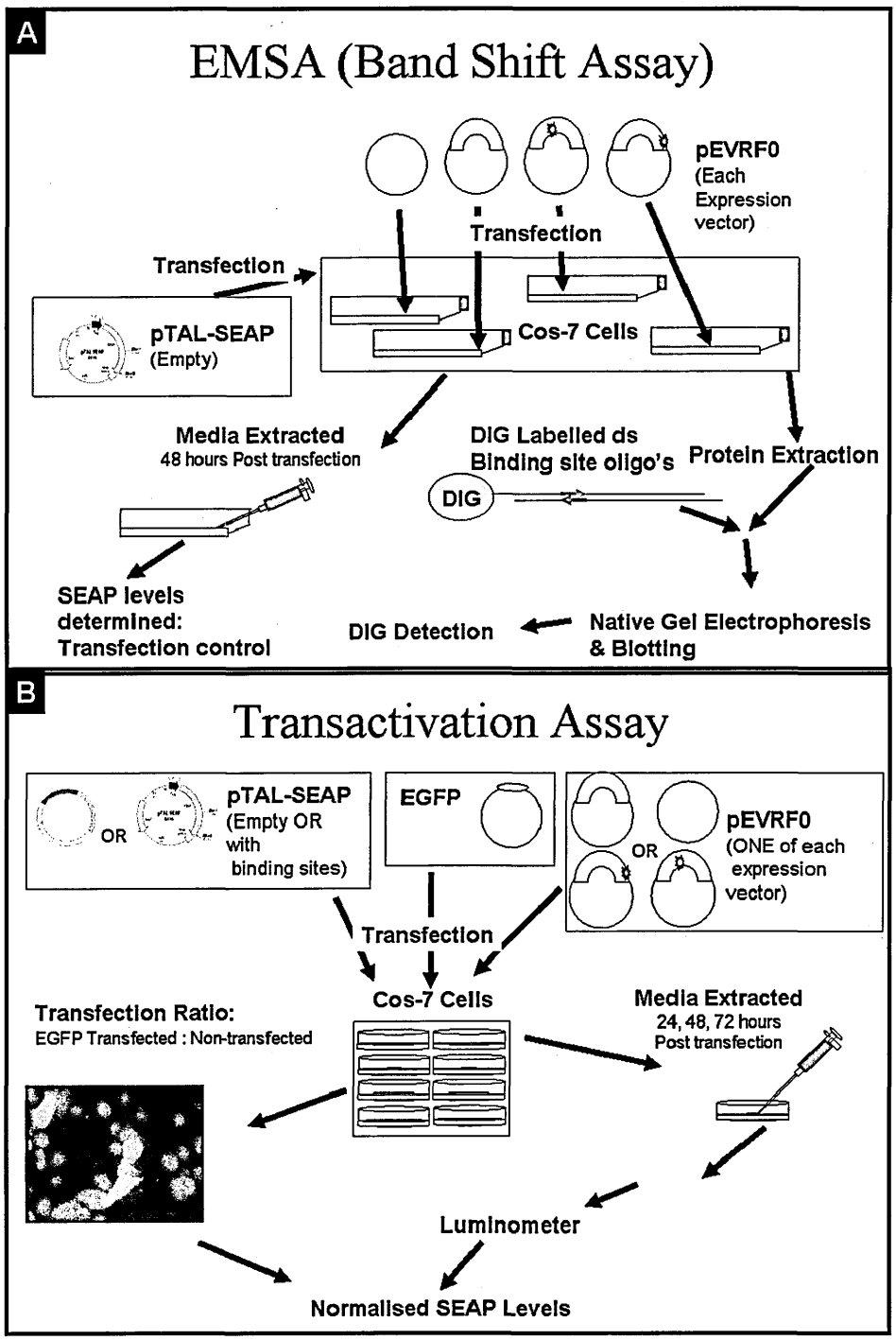


Figure 6.2 EMSA & Transactivation Assay schemes
Schemes planned for the study of the effects of the W174R and V410F mutations on the DNA binding affinity and transcriptional activity of FOXF2. A: For the EMSA, each of the pEVRF0 expression plasmids ($FOXF2^{W174R}$, $FOXF2^{V410F}$, $FOXF2^{PG}$ and the empty pEVRF0 vector) were transfected into separate Cos7 cultures alongside the empty pTAL-SEAP vector which acted as a transfection control. Media was extracted to measure alkaline phosphatase levels, which it was thought would be secreted at levels that varied only according to the transfection efficiency of the control vector in each sample. Nuclear protein was extracted from the cells and added to a DIG

labelled probe which contained the FOXF2 binding sequence. This mixture was then run on an electrophoretic gel in order to visualise the binding efficiency of each FOXF2 isotype. B: The transactivation assay involved the transfection in triplicate of each FOXF2 expression plasmid and a reporter. The reporter was pTAL-SEAP with multiple copies of the FOXF2 binding site adjacent to the pTAL promoter. pTAL-SEAP without the binding sequence was also added to some of the transfections as a control against which the regulation of SEAP could be measured. GFP was also added as a transfection control.

6.2.1 Mouse cDNA Expression Vector Construction

Foxf2 cDNA primers were designed to the 5' and 3' UTRs of the gene with the restriction sequences SalI and BamHI on separate primers and a random sequence tail at the 5' end of each primer. SalI and BamHI were selected as restriction sites because they weren't present in FOXF2, so digestion would not occur in the plasmid at extraneous positions. Both enzymes also function at 100% efficiency in a single buffer:- NEB3.

Construction of the plasmid was planned as follows:

1. RNA extraction from mouse tissue known to express Foxf2
2. Reverse transcription of RNA to produce cDNA
3. Amplification of cDNA with the primers designed to Foxf2 UTRs as shown in figure 6.3
4. SAll and BamHI restriction enzyme digestion of the expression plasmid.
5. Treat the linearised plasmid with phosphatase in order to remove phosphate groups to inhibit self ligation of the plasmid and to enable ligation of the amplified cDNA product.
6. Ligation of the cDNA product and plasmid.

7. Transformation of competent bacterial cells with the ligation product.
8. PCR of transformed colonies with primers designed to surround the multiple cloning site of pIRES-EGFP plasmid.
9. Size determination of PCR products.
10. Purification of plasmid from transformed bacterial cultures.
11. Sequence confirmation of insert presence and integrity.

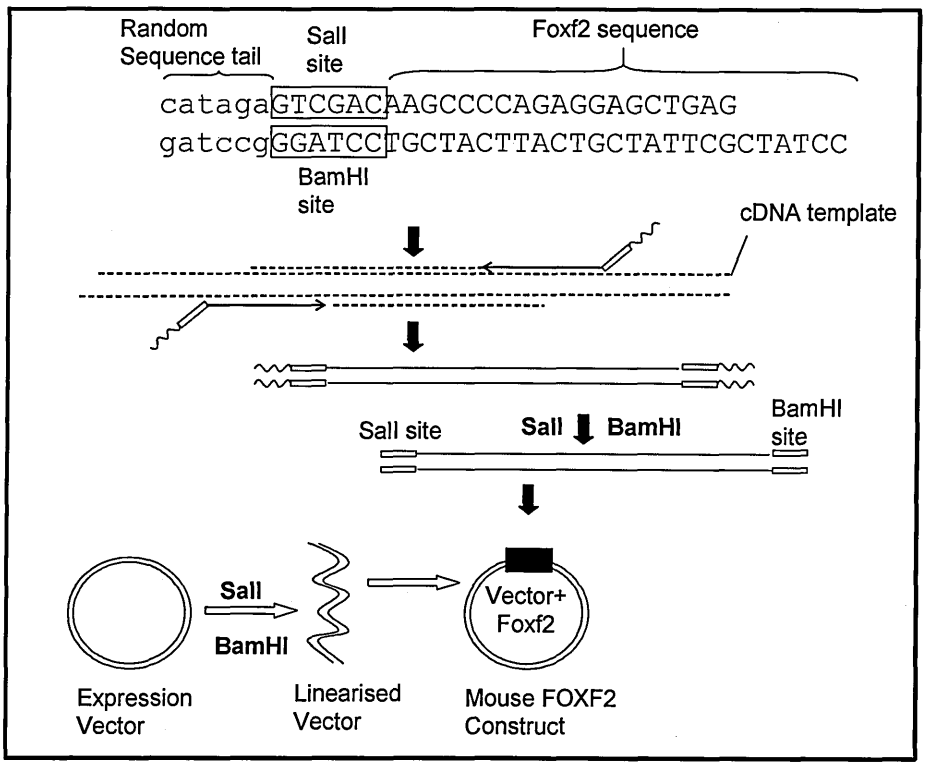


Figure 6.3: Intended Foxf2 Expression plasmid design
Approach for the Intended Construction of a Mouse Foxf2 expression construct. Primers were designed within the 5' and 3' UTR of mouse Foxf2. SalI and BamHI restriction site sequence were incorporated at the 5' end of the gene sequence of these primers for the later insertion into a pIRES-EGFP expression vector. A randomised 6bp tail was added to the restriction sequence to minimise the opportunity for inappropriate annealing and interference with Foxf2 amplification.

RNA was extracted from the lungs and eyes of ~8 week old mice as described in section 2.2.20. These tissues were selected because according to expression studies

that were carried out by Aitola and co-workers [188], the gene maintains expression into adulthood in these mouse organs.

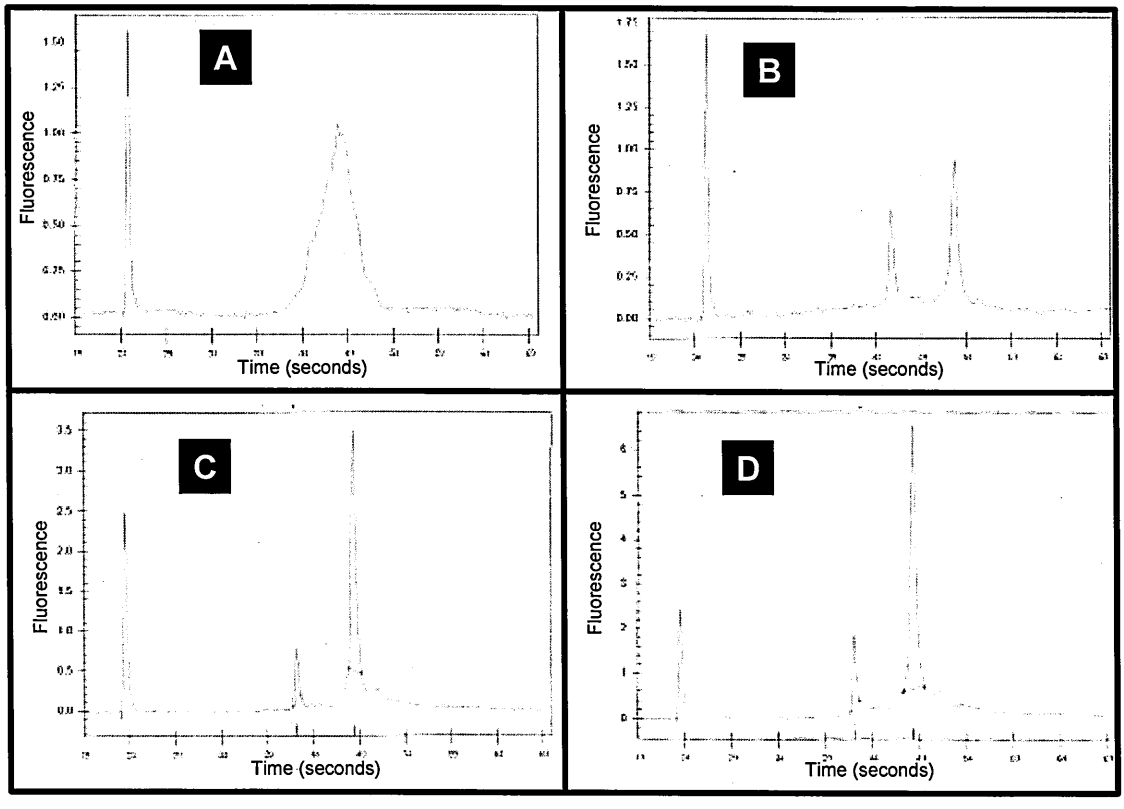


Figure 6.4: RNA purities
RNA Agilent Analysis for Eye and lung tissue from 1 male and 1 female heterozygote individual. Good quality, undegraded RNA produces two distinct peaks resulting from 18S and 28S rRNA in a 28S:18S ratio of ~2; a ratio which decreases as degradation proceeds in RNA samples. These two cytoplasmic ribosomal components together with the 5S ribosomal subunit constitute ~95% of total RNA and are therefore an easily detectable indicator of the general quality of total RNA. RNA from eye tissue of the female produced a broad, indistinct peak, suggesting that either degradation or contamination had occurred in this sample. B: RNA from eye tissue of the male produced sharp but low peaks – indicating a low yield (34.07ng/ul) of clean and good quality RNA. Total RNA from lung tissue from the female (C) from the male (D) were both of high quality, at concentrations of 23.13ng/ul and 53.99ng/ul respectively.

Initial attempts to obtain a cDNA product with the tailed primer set described in figure 6.3 were unsuccessful. However, following optimisation and changes to the Reverse transcription (RT) kit used, cDNA which produced a strong sharp PCR product from the primers Foxf2cDNAnested4 were obtained with the conditions

described in section 2.2.22. These products however were significantly shorter than the predicted size of 2340, so a range of different primer sets were applied to these RT products in order to obtain Foxf2 cDNA. None of the product sizes were of the correct size as shown in table 6.1. However, when 2 of the shorter products were sequenced, they showed exact sequence identity with the expected product for the 3' half of the molecule. The 5' half appeared truncated to different extents within each product, but in each case, sequence was missing from the product near the 5' primer.

Trehalose was added to some of the reverse transcription reactions because its presence allows enzymatic reactions to proceed at higher temperatures than is otherwise possible [273, 274]. This is useful when performing reverse transcription because higher temperatures can remove problematic secondary structure within RNA thus facilitating the synthesis of full length cDNA. I was only possible to amplify Foxf2 cDNA in the presence of trehalose.

Forward	Reverse	Expected Size	Observed size (Approx)
Foxf2cDNA4_nested	Foxf2cDNA4_nested	2340	1300
Foxf2cDNA3	Foxf2cDNA3	1728	700
cDNA5 F	Foxf2cDNA3	1475	700

Table 6.1: Expected and Observed Product sizes from Foxf2 cDNA PCRs.
These were the preliminary attempts to produce full length cDNA from Foxf2^{W174R} tissue. The product sizes were smaller than expected and did not therefore contain full length FOXF2 cDNA.

Following the initial failure to produce full length cDNA directly, another approach was taken. Primers were designed for the production of shorter overlapping fragments of the cDNA which it was envisaged would be easier to synthesise and could be joined together with restriction enzymes following amplification and purification of the products.

These primers however also failed to produce product of the expected size –each pair produced two smaller products as can be seen on table 6.2. Amplification of the shorter fragments of cDNA was intended to ease the production of full length cDNA. A series of restriction digests would have been required with this method. Both bands from each product were therefore sequenced. It was speculated that the bands could together contain the required sequence and although an unlikely option due to the technical challenges involved, one potential strategy was the utilisation of additional restriction digests to fuse all the required cDNA sequence.

BLAST Alignment of cDNA Fragment sequence with Ensembl v45 Foxf2 Transcript				
	Large PCR Product 1	Small PCR Product 1	Large PCR Product 2	Small PCR Product 2
Ensembl Sequence Range F (A)	630-824	-	1342-1713	-
Ensembl Sequence Range F (B)	311-352	-	826-863	-
Ensembl Sequence Range R	626-897	620-902	1550-1665	1560-1658
Region between F alignments in Fragment Sequence	94	Not Applicable	27	Not Applicable
Region between F alignments in Ensembl Transcript	275	Not Applicable	480	Not Applicable
Total readable aligned seq	314	283	410	99
Total fragment sequence	408	283	437	99
estimated size on gel	480	410	530	320
Expected Size	644		966	

Table 6.2: Incomplete Foxf2 cDNA fragments
Alignments of the sequence from cDNA PCR products produced from cDNA PCRs with the primers Foxf2cDNAsmlRi and Foxf2cDNA5F (Small and Large PCR product1); and with the primers Foxf2cDNA3_R and Foxf2cDNAsml_F (Small and Large PCR product 2). Two products resulted from each PCR. The Forward reads that produced viable sequence had 2 separate regions of alignment, these have been nominated as A and B. Forward sequencing of both small fragments failed to produce viable data, indicating that these fragments were missing the 5' end of the required sequence.

Despite these unexpected product sizes, sequence data suggests that they are of Foxf2 origin. All of the amplified Foxf2 primer sets produced a strong product and sequence from these fragments aligned well with the sequence of Foxf2 in Ensembl, but each with a section of the 5' end missing. In each case a section of sequence data was missing from close to the 5' primer on sequencing the products.

The amount of total aligned sequence does not account for the entire PCR product in any of these samples. However, information about the nature of the PCR products can be obtained from the available sequence data.

A BLAST search of the PCR product sequence produces significant alignment uniquely with Foxf2, suggesting that these are PCR products of Foxf2 cDNA. In both large PCR products, the initial 40-50bp that immediately follows each 5' primer, aligns with the Ensembl transcript. A larger region of alignment at the predicted 3' end of the expected sequence is also seen in both PCR products. However, between these regions in both PCR products, lies sequence with no alignment matches to either Foxf2 or any other genomic sequence. The region between alignments is significantly larger in the Ensembl sequence than in either PCR product sequence (94bp and 27bp for the cDNA sequence of PCR products 1 and 2 respectively compared to 275bp and 480bp for the Ensembl equivalent regions of PCR products 1 and 2 respectively)

This phenomenon can only be seen in the forward sequencing of each PCR product, the reverse primer does not produce the sequence seen with the forward primers in the 5' aligned region, probably due the distance from the 5' region to the reverse primer sequence.

Each PCR also produced a second, smaller band which was also sequenced. Again, when analysed with BLAST, these PCR product sequences only produced significant alignment with Foxf2 and no other genomic sequence. The forward sequencing primer failed to produce any sequence for either PCR product, despite attaining sequence of high quality for the reverse primers. This suggests that the 5' end of the expected PCR product is missing from these smaller amplicons.

The discovery of these unusual larger PCR products was a by-product of an attempt to produce a full length mouse cDNA for the manufacture of an expression plasmid. Because the cDNA was proving difficult to obtain, it was decided that an available human FOXF2 expression construct would be used instead. Further characterisation of the cDNA was therefore not considered a priority but an investigation into whether these are genuine transcriptional isoforms and if so, the identification of the precise breakpoints could be interesting, particularly as there is only one intron in the gene.

Forward Primer	Reverse Primer	Product Size Calculated from Sequence	Product Size Estimation from Agarose Gel
Foxf2cDNA4nested	Foxf2cDNA4nested	2340	1300
Foxf2cDNA3	Foxf2cDNA3	1728	700
cDNA5 F	Foxf2cDNA3	1475	700
cDNA5 F	Foxf2cDNAsmlRi	644	500
Foxf2cDNAsml F	Foxf2cDNA3	966	520 & 300
Region of Primer Overlap		-	-

AGCTGTCCAGCTCTGGGCAGCGTCGCCTCAGTGGCTTAGGAGACGGTTGCGCAAGGAGCCGGCCGGAGGCTCGGA
AGAGGGATGCGCGGGGCGTTGCCTCCGACCCGCCGCCGCCGCCGCCGCCCGAAGCCCCAGAGGAGCTGAGGGAGG
CGACGCCGAAGCGCTGGCCCGCAGTGGCCCGGGCTGCAGCGCGGCCGGCGCAGTAGGGCACTCGCCCGATTTGTG
GACCGACCCGGCTCCGGCGCCGCCGATCCCGGTCCGGGGTCCGCCCCCCAGGCCCGGCCTCGTTCCCGATTCCCA
GATGAGCACCGAGGGCGGGCCTCCGCCACCCCCGCCGCGCCCGCCGCCTGCCCCACTCCGCCGCGCGTGCAGCCC
GGCGCCCGGCGCGCTCCAGGCCGCCTTGATGAGCCCGCCACCCGCCGCCACCCTGGAGTCCACTTCGTCGTCGTC
ATCATCATCCTCTGCCTCCTGTGCCTCGTCCTCTTCTAACTCCGTCAGCGCCTCGGCCGGTGCTTGCAAGAGTGC
GGCTAGCAGCGGCGGCGCGGGCGCCGGGAGTGGAGGCACCAAGAAGGCAACCTCGGGGCTGCGGCGGCCGGAGAA
GCCTCCTTACTCGTACATCGCGCTCATCGTCATGGCCATCCAGAGCTCGCCCAGCAAGCGCCTGACGCTCAGTGA
GATCTACCAGTTCCTACAGGCGCGCTTCCCCTTTTTCCGTGGCGCCTACCAGGGCTGGAAGAACTCCGTGCGCCA
CAACCTCTCGCTCAACGAGTGCTTCATCAAGCTGCCCAAGGGCCTCGGGAGACCTGGTAACGGCCACTACTGGAC
CATCGACCCGGCCAGCGAATTCATGTTTGAGGAGGGTTCGTTCCGCCGCCGGCCGCGCGGCTTCAGGCGGAAGTG
CCAGGCTCTCAAACCCATGTACCATCGCGTGGTGAGCGGCTTGGGCTTCGGGGCCTCGCTGCTGCCCCAGGGCTT
CGACTTCCAAGCGCCCCCGTCGGCGCCTCTGGGTTGCCACGGTCAAGGCGGTTATGGTGGCCTCGACATGATGCC
CGCGGGCTATGATACAGGGGCGGGTGCTCCGGGCCACGCGCATCCACACCACCTCCACCACCACCACGTCCCCCA
CATGTCGCCCAACCCGGGCTCCACCTATATGGCCAGCTGCCCGGTGCCCGCAGGTCCTGCGGGCGTCGGTGCGGC
AGCGGGTGGCGGCGGTGGCGGCGGGGACTATGGGCCGGACAGCAGCAGCAGCCCTGTGCCCTCATCCCCGGCTAT
GGCAAGCGCCATTGAGTGTCACTCGCCCTACACTAGCCCTGCGGCACATTGGAGCTCGCCTGGCGCTTCACCTTA
CCTCAAGCAGCCGCCTGCCCTGACGCCAAGCAGTAATCCCGCGGCCTCTGCTGGTCTGCACCCCAGCATGTCTTC
CTACTCGTTGGAGCAGAGCTACTTGCACCAGAACGCCCGCGAGGATCTCTCAGTCGGACTGCCCCGTTACCAGCA
CCACTCCACTCCAGTGTGCGACAGGAAAGATTTCGTCCTCAATTTCAATGGCATTTCTTCTTTCCACCCTTCCGC
TAGTGGCTCTTATTATCACCATCATCACCAGAGCGTGTGCCAAGATATTAAGCCCTGTGTTATGTGAATGGACAG
AGGCCTTGAAGGCCACTCTGCTCTCCCTCTTCTGCCTTCTTCTCCTCTCCCCTCCGAGAGGGGCAGCCAGGAACT
GCAACGGACTCACACTGTGCACGCTGGATAGCGAATAGCAGTAAGTAGCACACCCATCACTTAGACAAATACCCA
AGGGAGTTCTGCTCACCGATATTTGCCCGGCCCCTGGAAGAGGAAACCTTTCGAAAGCTAATATCCCAGAAGAGC
GACAGACAGAGGAGGTGACTACATGTAAGACATATGTTACTGTGTGGAGGACATAAAACTTTTCAGTTCTGGGGT
GGCCATTGCATTCACTAATCAGGGTCTGAAAAGGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGCC
TTTTCAAAATTGCAGTGCTAAAAGCACACAATTTCAAGAAAGCCTTCTCTAGTTCCCTGGCTCAGTAGGACATTT
CTTCCACCCACTCCCAGTCCCCCAACACTCCCTTCTTCGGATACAGGTGCCAAAGAACATTGTGAAGAAATGAAG
AACCGATAGTCTAGTTTAAGAAAGTGGCTCTCAGTATTGTGACAGTACATTTTTACAAGGTTGTTTTTCTACCAC
CGTATTTTTAAAGTATTTTTATGATCTTCGTATACTCACACTTCGCTTGTATTGTAAAAGGAGGGTATATTTGCA
CTTATGTATACTTTGACAGTTTGCCAAGATATTTTGATGTAAGGTTTTTTTTTTTTCAATAAAATGTGTATAAC

Figure 6.5: cDNA fragments in relation to the transcript
The Foxf2 transcript as predicted by Ensembl v45 and positions within the sequence of the primers used to amplify cDNA. Underlined, bold sequence is translated into protein.

6.2.2 A Human FOXF2 Expression Construct

A FOXF2 expression construct was kindly supplied by Dr Cheryl Gregory Evans. This was originally constructed by Marika Hellqvist and co-workers [203] who inserted a FOXF2 cDNA sequence into a pEVRF0 plasmid [204]. The inserted cDNA is not the full length and the resulting protein is therefore missing 39 amino acids from the amino terminus.

The pEV vector set of which pEVRF0 is one member, allows the expression of incomplete coding sequence in any reading frame by selection of a specific vector according to the position of the first coding base available. Three of the four vectors within the set are designed for inserts missing their 5' end, each of these translates sequence within a different reading frame. The translation initiation region of HSV TK (Herpes Simplex Virus - Thymidine Kinase) is located upstream of a multiple cloning site at different positions in each vector to allow for different reading frame insertions [204]. The translated products are therefore fusion proteins that contain 9 N-terminal amino acids of vector origin. The promoter region is derived from the human CMV (Cytomegalovirus) promoter and the origin of replication is from the SV40 (Simian Virus 40) *ori* locus.

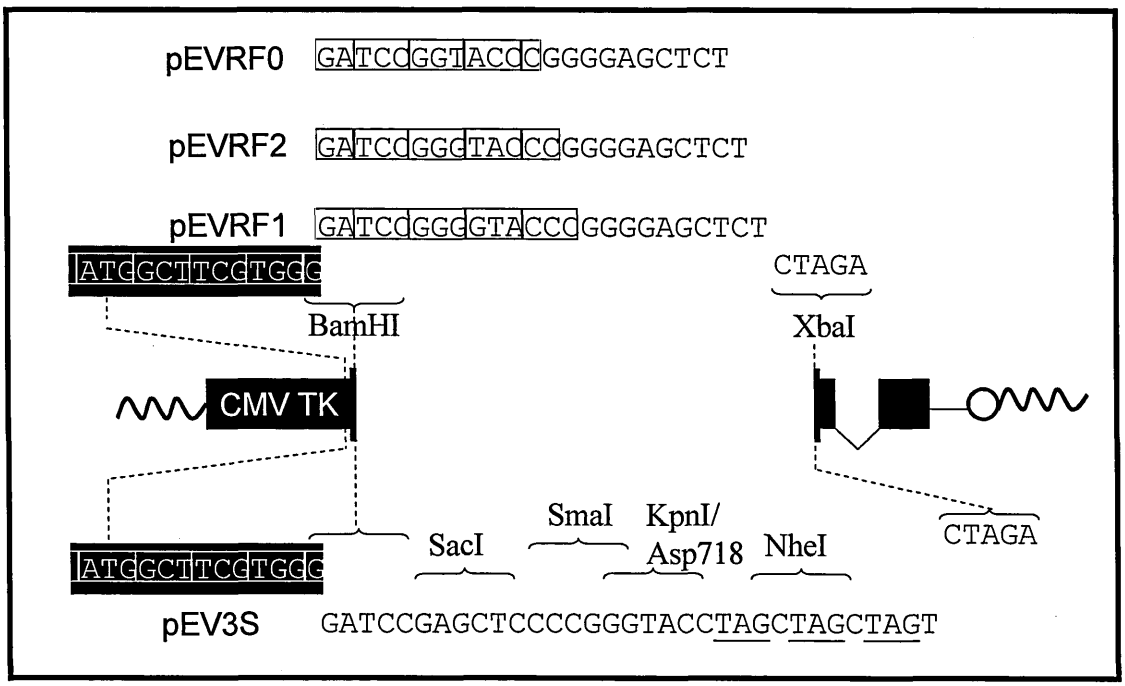


Figure6.6: pEVRF0 Expression plasmid
The pEV plasmid set, copied from Matthias et al [204]. pEVRF0, 1 and 2 allow the expression of products which are missing their 5' initiation region by the presence of a Thymidine kinase initiation sequence adjacent to the cloning site. The multiple cloning site allows for insertions within each possible reading frame

In addition to the protein truncation and presence of amino acids encoded by the vector, three amino acids which Hellqvist identified as the 3 terminal FOXF2 residues of the insert differ from the Ensembl FOXF2 sequence; these are Q F R in the construct whereas the sequence in Ensembl identifies the amino acids at this position as: A A A within a (9X) poly-alanine repeat. The predicted product from this construct is therefore either a fusion of 12 foreign amino acids onto a truncated FOXF2 that is missing 39 amino acids if the Ensembl cDNA sequence is correct, or a 9 residue product fused to a FOXF2 protein that is missing 36 amino acids – if the cDNA and associated peptide sequence described in Hellqvist's paper [203] is accurate.

Cheryl Gregory-Evans and Helena Vieira had subsequently mutagenised this construct to include an additional glycine within a polyglycine tract (Gly302-306ins3bp), replicating an insertion mutation which they had identified in a family of glaucoma patients [155].

The two mutations that were identified in the ENU archive which required further analysis were W174R and V412F in mouse Foxf2. In the human gene, the designation of the tryptophan to arginine mutation remains unchanged at amino acid position 174, but the valine position is at amino acid position 410, not 412 as is the case in mouse. The valine to phenylalanine mutation will therefore be referred to as V410F for all subsequent work involving the human FOXF2 plasmid.

6.2.3 Vector Mutagenesis

The difficulties experienced with the Foxf2 cDNA production meant that another approach was required for the production of an expression plasmid. A human FOXF2 plasmid was available which could be mutagenised by site-directed mutagenesis to induce the equivalent mutations that were obtained from the ENU mouse archive. Although the 5' end of this vector was missing, the entire forkhead domain and all the protein sequence 3' of the forkhead domain was present.

6.2. 4 Mutant Plasmid Production

Site directed mutagenesis was used to introduce the mutations of interest into the human FOXF2 expression plasmid. This procedure required the design of overlapping primers that contained the target sequence with the desired mutation as described in section 2.2.30 (see also figure 6.7). These primers were used to amplify the entire plasmid using a template plasmid that had been harvested from Dam+ host bacteria. DpnI is added to the resulting product which degrades only Dam+ DNA so should remove the non mutagenised parental plasmid and leave the mutagenised PCR product. This can then be transformed into a host bacteria to repair the nicks for the finished mutagenised plasmid.

The polyglycine mutant was amplified using the conditions described in section 2.2.5, programme 4. This was followed by DpnI digestion with 10 units of DpnI in 50ul PCR product mixture. Transformation of XL1-Blue cells was performed as described in section 2.2.18 with 1ul, 2ul and 5ul volumes the DpnI reaction mixture.

14 colonies were grown overnight in LB with ampicillin. 750ul of these cultures were stored as glycerol stock and the remaining culture purified using a Wizard prep kit as described in section 2.2.1 before direct sequencing with primers designed to the region that contained the site of the desired mutation (see table 6.4 for all primers that were used to sequence potential mutants). Of these purified vectors, 7 were missing sections of the expected sequence. One of the seven remaining vectors was successfully mutagenised to remove the additional GGC triplet in order to produce the wildtype version of the FOXF2 plasmid.

This procedure was then repeated to induce the W174R mutation and the V410F mutation with the primer sets and amplification programmes as described in table 6.3.

Primers For Site Directed Mutagenesis				
Mutation	**Primer set**	**Forward**	**Reverse**	**PCR Programme**
Wildtype from Gly302-306ins3bp	HFOXF2_7G-6G	GGGGGCGGCGGCGGCGGCG ACTACGGGCCGGACAGC	GCTGTCCGGCCCGTAGTCGC CGCCGCCGCCGCCCCC	55C annealing. 68C extension. 16 cycles
W174R	HFOXF2_W-Ri	CGGCAAGGGCCACTACAGGA CCATCGACCCG	CGGGTCGATGGTCCTGTAGT GGCCCTTGCCG	58.8C annealing. 72Cextension. Betaine. 12 cycles
V410F	FOXF2_V-F	GTGTGTGACAGAAAAGATTTC TTCCTCAACTTCAATGGG	CCCATTGAAGTTGAGGAAGAA ATCTTTTCTGTCACACAC	66.8C annealing. 72C extension. Betaine. 12 cycles

Table 6.3: Site directed mutagenesis thermal cycling conditions
Primers and PCR conditions which produced the required mutations within the FOXF2 pEVRF0 plasmid.

Following sequencing identification of the mutant plasmids, further sequencing with the primer sets listed below was performed to ensure that polymerase induced errors had not occurred during amplification, and that the FOXF2 open reading frames were maintained.

Primer Sets Used for Plasmid Integrity Testing

Primer ID	Forward Primer	Reverse Primer
pEVRF0: N	GGGGGATCTTGGTGGCGTG	Not Applicable
HFOXF2_1A	Not Applicable	GGATGGCCATGACGATGAGC
HFOXF2_1B	AGCCGCCCTACTCGTACATC	ATGTCGAGGCCGCCGTAG
HFOXF2_1C	CAGGGCTTCGACTTCCAG	TGCAAGTAGCTCTGCTCCAG
HFOXF2_2A	AATGCCACTCGCCCTACAC	Not Applicable
pEVRF0: C	Not Applicable	CCCTGAAAACTTTGCCCCCTCC

Table 6.4: Primers for Testing the Integrity of pEVRF0 Constructs
Primer sets used for expression plasmid integrity testing. The pEVRF0: N & C primers are designed to the vector sequence surrounding the insert and the HFOXF2 primers are designed to human FOXF2 sequence. The primers at the terminal ends of the expressed sequence were not included for this screen because the sequence homologous to HFOXF2_1A (forward) is missing from the truncated sequence of the plasmid and the sequencing data from HFOXF2_2A (reverse) would be redundant due to the pEVRF0_C sequencing data. The Vector sequence primers were used to check for construct integrity in addition to the FOXF2 primers because they provide sequence data for the region which spans the insert boundaries.

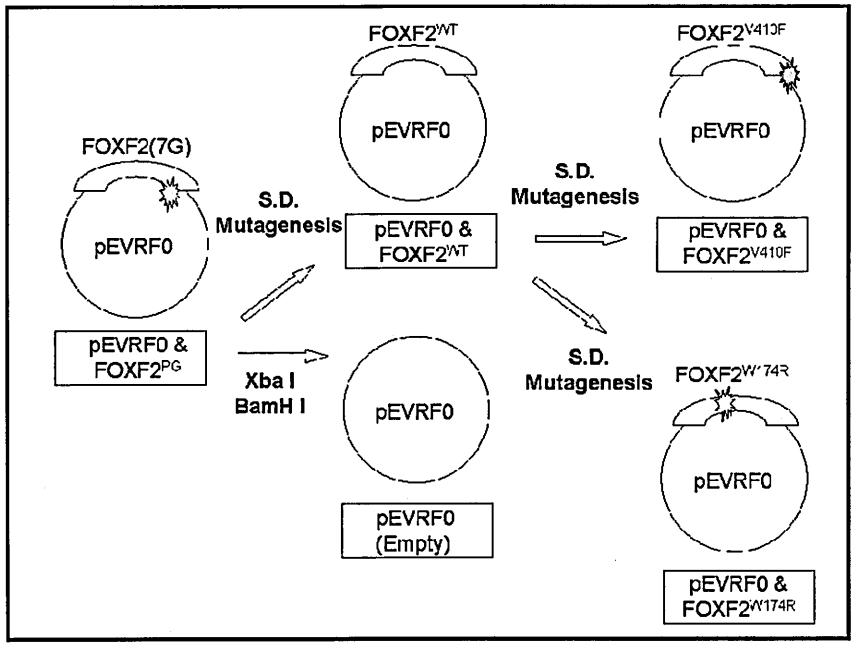


Figure 6.7: FOXF2 Expression Construct Mutagenesis
Mutagenesis strategy for the FOXF2 expression constructs. The construct supplied contained the polyglycine expansion that was seen in human glaucoma carriers. This mutation was removed by mutagenising the polyglycine region with wildtype primers before further site directed mutagenesis was performed to induce the V410F mutation and W174R mutation. The empty vector was produced with a double digest of the original construct using XbaI and BamHI.

A: Human FOXF2 Protein Sequence

MTTEGGPPPAPLRRACSPVPGALQAALMSPPPAAAAAAAAAPETTSSSSSSSSASCASS

SSSSNSASAPSAACKSAGGGGAGAGSGGAKKASSGLRRPEKPPYSYIALIVMAIQSSPS

KRLTLSEIYQFLQARFPFFRGAYQGWKNSVRHNLSLNECFIKLPKGLGRPGKGHYWTID

PASEFMFEEGSFRRRPRGFRRKCQALKPMYHRVVSGLGFGASLLPQGFDFQAPPSAPLG

CHSQGGYGGLDMMPAGYDAGAGAPSHAHPHHHHHHHVPHMSPNPGSTYMASCPVPAGPG

GVGAAGGGGGGDYGPDSSSSPVPSSPAMASAIECHSPYTSPAAHWSSPGASPYLKQPPA

LTPSSNPAASAGLHSSMSSYSLEQSYLHQNAREDLSVGLPRYQHHSTPVCDRKDFVLNF

NGISSFHPSASGSYYHHHHQSVCQDIKPCVM

.....

B: FOXF2 & pEVRF0 construct:

----------MASWGSGTQFRAAPETTSSSSSSSSASCA

SSSSSSNSASAPSAACKSAGGGGAGAGSGGAKKASSGLRRPEKPPYSYIALIVMAIQSS

PSKRLTLSEIYQFLQARFPFFRGAYQGWKNSVRHNLSLNECFIKLPKGLGRPGKGHYWT

IDPASEFMFEEGSFRRRPRGFRRKCQALKPMYHRVVSGLGFGASLLPQGFDFQAPPSAP

LGCHSQGGYGGLDMMPAGYDAGAGAPSHAHPHHHHHHHVPHMSPNPGSTYMASCPVPAG

PGGVGAAGGGGGGGDYGPDSSSSPVPSSPAMASAIECHSPYTSPAAHWSSPGASPYLKQ

PPALTPSSNPAASAGLHSSMSSYSLEQSYLHQNAREDLSVGLPRYQHHSTPVCDRKDFV

LNFNGISSFHPSASGSYYHHHHQSVCQDIKPCVM

Figure 6.8 Amino acid sequence of FOXF2 and the truncated FOXF2 of the peptide

Amino acid sequence of FOXF2 and the truncated FOXF2 of the peptide

Amino Acid sequence of A: Human FOXF2 and B: The truncated FOXF2 fusion product produced by the pEVRF0 construct. Red letters indicate the position of the forkhead domain within each molecule. The vector encoded sequence is underlined on the fusion product . The polyglycine expansion region is highlighted in yellow and shows the human wildtype copy number (6 residues) in A, and the expanded copy number (7 residues) in B. The first 3 amino acids of FOXF2 origin in sequence B (QFR) are not present in sequence A which was sourced from Ensembl but were identified as part of a cDNA derived peptide in Hellqvist's paper [203]. Blue underlined text is amino acid sequence originating from the plasmid.

6.2.5 Production of the Empty pEVRF0 Expression Plasmid

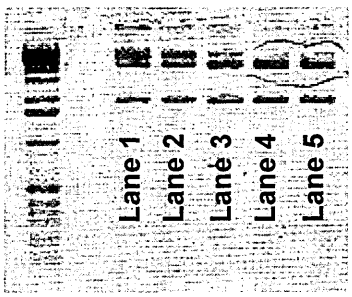


Figure 6.9 Evidence of Empty pEVRF0 plasmid
A range of XbaI and BamHI concentrations were added to 2ul (approx 400ng DNA) of plasmid (following purification with a wizard kit) to ensure its complete digestion. Lane 1 = Original digest; Lane 2 = Original digest then a further 10 units of BamHI; Lane 3 = Original digest + 10u XbaI; Lane 4 = Original digest + 10u BamHI + 10u XbaI; Lane 5 = Original digest + 20u BamHI + 20u XbaI. Complete digestion of the linearised plasmid has occurred in lanes 4 and 5 where only the cut plasmid (middle band) and free insert (lower band) is present, and not the upper band (undigested plasmid + insert). Double digests were possible because both enzymes are 100% efficient in buffer 2 in the presence of BSA

Following digestion with XbaI and BamHI and subsequent DNA purification with a Qiaquick PCR purification kit (Qiagen), the sticky ends of the linearised pEVRF0 vector were filled using DNA polymerase I (klenow), before recircularising with blunt ended ligation as described in sections 2.2.17 and 2.2.30.

6.2.6 EMSA Results

Two potential oligonucleotide probes were labelled with Digoxigenin-11-ddUTP (DIG) – these probes are oligonucleotides (described in section 2.2.31) that contain the FOXF2 binding site and would therefore be expected to allow binding of the FOXF2 protein. Serial dilutions of these DIG labelled oligonucleotides and a control oligonucleotide that was labelled under standard conditions by the manufacturer of the labelling kit (sequence from the lac Z region of pUC and M13 plasmids: pTTG GGT AAC GCC AGG GTT TTC CCA GTC ACG). In order to visualise the efficiency of labelling, the membranes were treated with an alkaline phosphatase (AP) conjugated anti-DIG antibody, before the application of the chemiluminescent AP substrate CSPD and subsequent exposure to X-ray film. Oligonucleotide concentrations for optimal signal strength and resolution were determined from the image produced (figure 6.10). Subsequent binding assays utilised the BglII oligonucleotide annealed to an unlabelled version of the MluI oligonucleotide at 50fmol/µl.

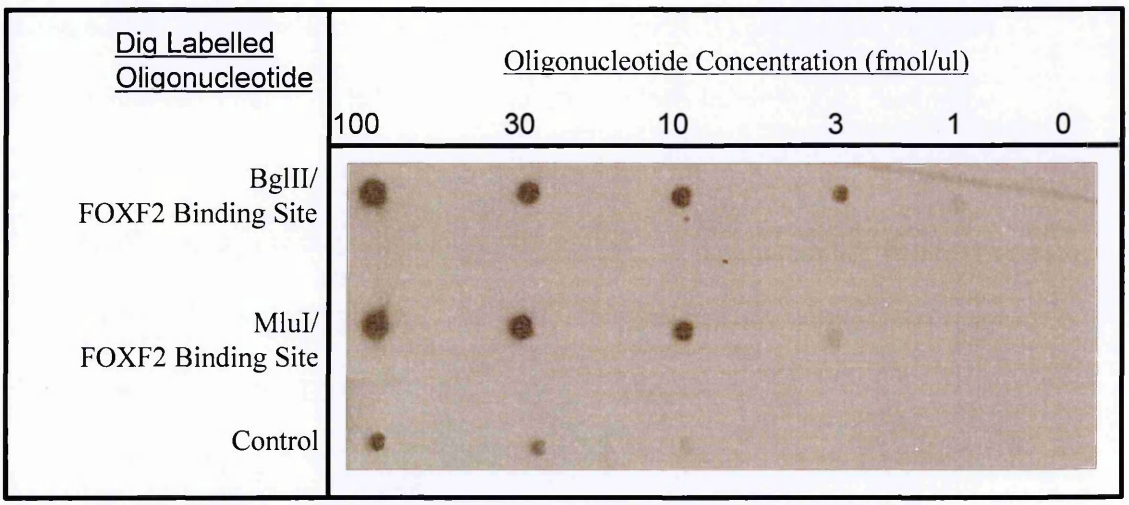


Figure 6.10: DIG oligonucleotide labelling efficiency tests
Following the DIG labelling procedure, the labelled oligonucleotides were serially diluted and blotted onto Hybond N+ membrane before cross-linking before detection of DIG levels with an alkaline phosphatase conjugated anti-DIG antibody.

Proteins that bind to DNA irrespective of the nucleotide sequence will be present within the experimental nuclear extract. This is likely to have a confounding effect on the interpretation of any analysis of specific binding of extracted proteins to the DIG labelled probe. The most common way to reduce this phenomenon is the addition of non-specific competitors. Different combinations of Salmon sperm DNA, CotI DNA and poly dI-dC were therefore tested to discover which competitor provided the greatest improvement to band quality. The non specific competitors that allowed the greatest clarity and resolution of visible bands was a combination of CotI and poly dI-dC as can be seen in figure 6.11.

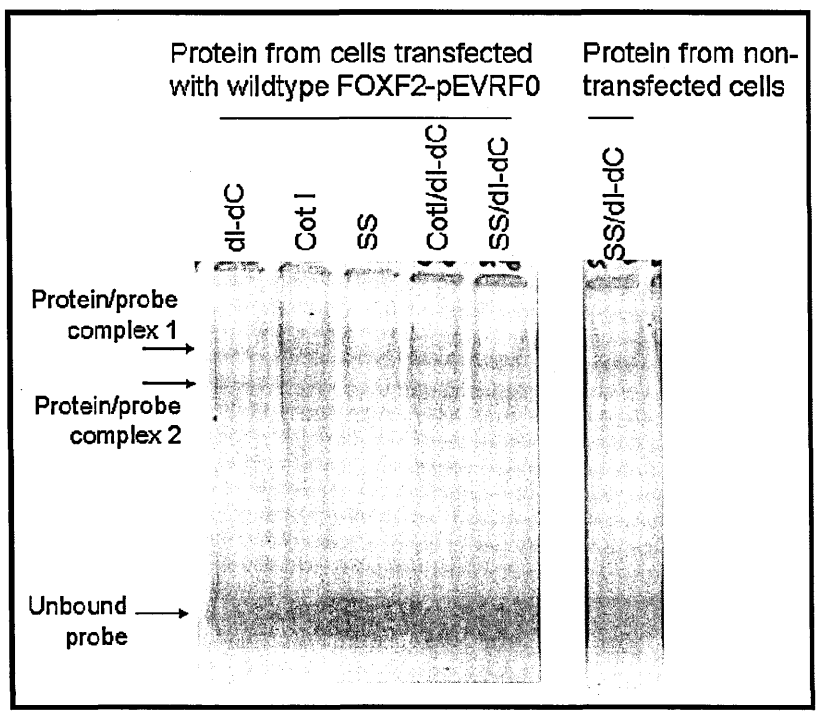


Figure 6.11: Non Specific Competitor Selection
This test was carried out with a range of different non specific inhibitors with 15ug nuclear protein extracted from cells that were transfected with the wildtype FOXF2 construct. The non-specific competitor which produced the strongest bands that still maintained a clear resolution was a combination of CotI and poly dI-dC. SS=salmon sperm DNA at 50ng/ul. dI-dC= poly dI-dC at 0.01U/ul. CotI = Mouse CotI DNA at 50ng/ul. A total of 2ul non-specific inhibitor was added to each sample.

Levels of nuclear extract within the binding mixture also required optimisation, so following an evaluation of protein concentration within the extracts, a range of different dilutions of the extract were added to the binding mixture. In an attempt to gain further insight into which if any of the visible bands were FOXF2/probe complexes, an anti-FOXF2 antibody was added to replicate samples. The presence of a supershift would have provided confirmation of the presence of FOXF2 in the nuclear extracts of transfected cells, but no supershift was visible (figure 6.12).

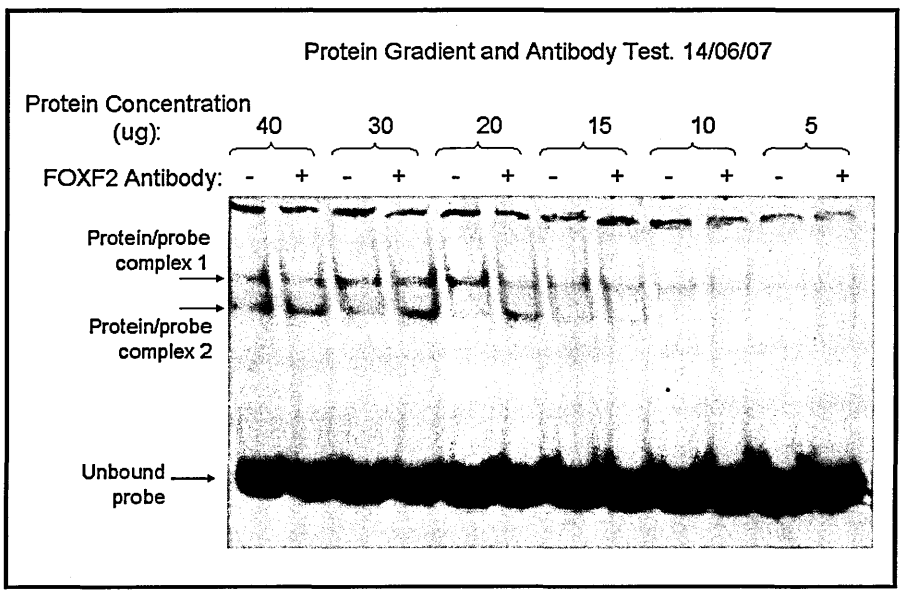


Figure 6.12: Protein concentration optimisation
Six different protein concentrations were tested with and without FOXF2 antibody 1 (Abnova). A supershift was not immediately apparent. Two potential protein/complex bands were seen when nuclear protein from the wildtype FOXF2 transfected cells were run with the probe in an EMSA in either the presence or absence of antibody. However, the samples containing antibody have greater signal strengths on the lower of the two prominent bands than do the samples without antibody. The apparent detection limit of the two probe-protein complex bands is between 10ug and 15ug protein.

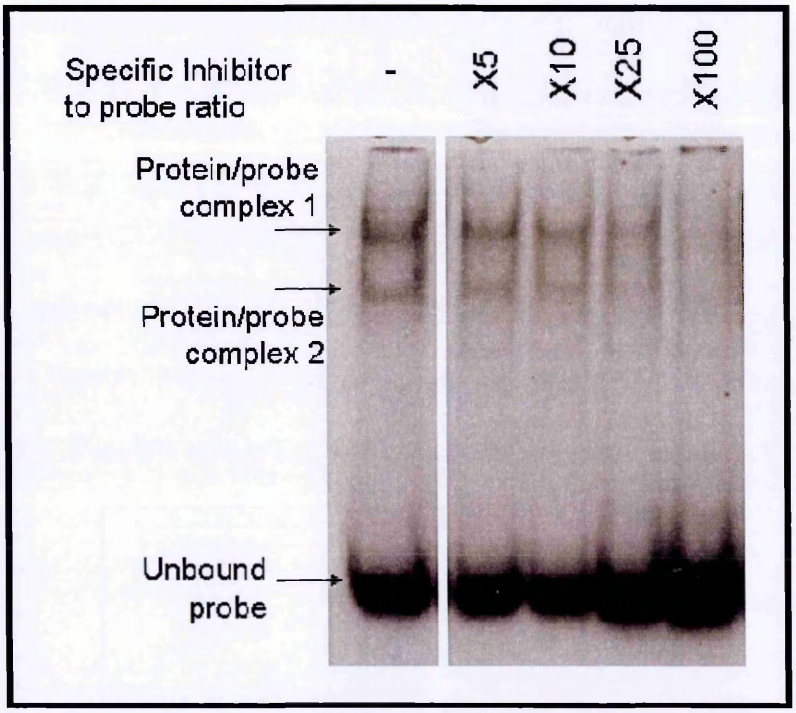


Figure 6.13 Specific competitor assay
Specific competitor assay. A reduction in the signal strength can be observed from both the upper and lower protein/probe complex bands corresponding with an increasing concentration of the specific inhibitor. This suggests that the specific inhibitor interferes with the binding of both bands and could lead to the conclusion that both bands may be the products of different forkhead genes. An increase in the signal strength from the unbound free oligonucleotide can also be seen to correlate with increasing specific inhibitor strength, reflecting the increasing amount of DIG labelled oligonucleotide that was unable to bind to a recognition site.

An additional approach taken to improve the clarity of the visible bands and to determine whether or not the binding interaction was specific - was the addition of a specifc competitor. A specific competitor in a band shift assay is unlabelled DNA that is identical in its sequence to the labelled probe. Any molecule that binds specifically to the probe is likely to have an equal affinity to the specific competitor and therefore an increase in the concentration of the competitor should result in an equivalent reduction in the amount of bound probe and therefore also in the strength of signal produced by the specific interaction. There is an inverse relationship between the strength of signal and the concentration of specific competitor in the

two prominent bands that can be seen in samples containing nuclear extract of FOXF2 transfected Cos-7 cells (figure 6.13 and 6.14). This result suggests that binding to both bands is specific. A supershift assay that was performed alongside another specific competitor test produced additional bands that emigrated slower than the suspected FOXF2 band (protein/probe 2 on figure 6.16). This would be expected if the probe was bound to FOXF2 given that a FOXF2 antibody should bind to FOXF2 thus increasing the size and slowing the progress of a DNA/FOXF2 complex. However, another band of greater intensity that emigrated faster than the potential FOXF2/probe band was evidence that the antibody may also be binding to something other than FOXF2. Further evidence against specific binding of the antibody was seen when the DIG probe was added to wells without any nuclear extract and a similar although weaker run profile was observed (figure 6.16 *) although this may have been cased by contamination from adjacent wells.

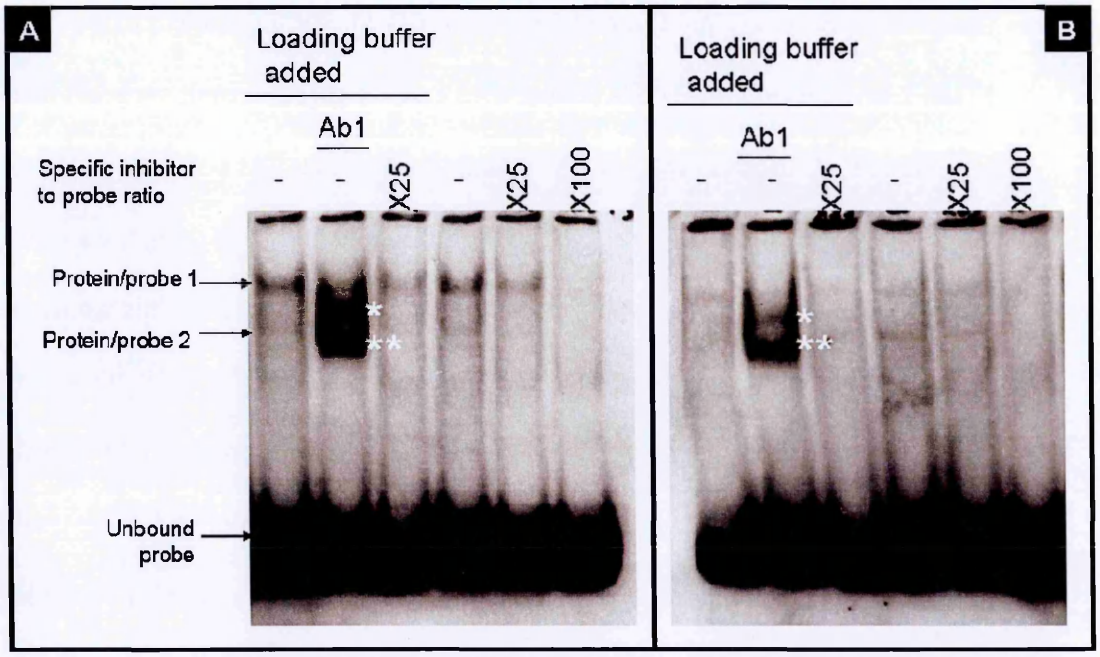


Figure 6.14: Specific competitor assay
A = 1 ul Cot I, 1ul dI-dC, 15ug WT protein. B = 1 ul Cot I, 3ul dI-dC, 30ug WT protein. Controls added: "Ab1" = Foxf2 Antibody 1 (Abnova) added. "X25", "X100" = Unlabelled probe (specific inhibitor) added at a concentration of X25 OR X100 the concentration of labelled probe. "-" = no specific inhibitor added. A supershift assay was performed alongside a specific competitor test in the hope that it may provide additional evidence as to the identity of a FOXF2 band. The addition of the antibody 1 (Ab1) appears to have enhanced the strength of the signal generated. * = potential supershift bands which are not present in any of the other samples that are higher than the suspected FOXF2 band (Protein/probe 2). ** = a more intense band below the suspected FOXF2 position. If non-specific binding to the antibody has occurred, it is possible that the lower band is the antibody bound to the probe. Another possibility is that the FOXF2 band was previously mis-identified and the lower band in lane 2 is FOXF2 bound to the probe and the higher band (*) remains the supershift. The reduction of signal strength caused by the presence of the specific inhibitor however, goes some way to discount this theory. There is also a noticeable weakening of the signal from the upper protein/probe 1 band. This band is not visible at all in the antibody lane in image 6.14B.

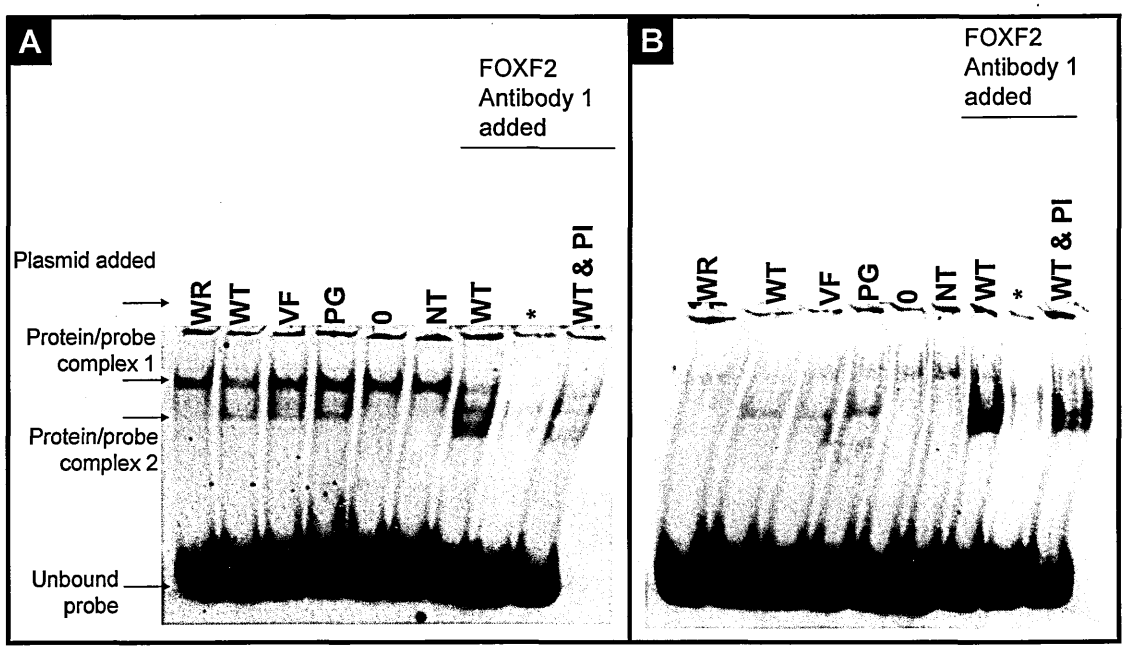


Figure 6.0-3 EMSA

Figure 6.15: EMSA of Cos-7 Cell lysate Following Transfection with FOXF2 Constructs
EMSA of FOXF2 with DIG labelled dsDNA containing the FOXF2 recognition sequence CAAACGTAAACAATCC. Nuclear protein was extracted from Cos-7 cells following transfection with a range of FOXF2 expression constructs. Controls are to the right of each gel image and include: 0 = nuclear protein extract from empty pEVRF0 vector; NT = nuclear protein extract from non-transfected cells. Antibody was added to extract from cells transfected with the wildtype FOXF2 expression plasmid (WT). Other samples tested were the $FOXF2^{V410F}$ construct (VF), the polyglycine mutant FOXF2 construct (PG) and the $FOXF2^{W174R}$ construct (WR). It was also added to the DIG probe alone (*) in order to determine whether a banding pattern occurred without the presence of protein. A Proteinase inhibitor cocktail (PI) was added to a mixture of nuclear extract from cells transfected with the wildtype expression plasmid and FOXF2 antibody 1 (Abnova). It was hoped that the presence of PI would reduce the potential for degradation of FOXF2 to impact on the quality of data – There is a more defined separation of the bands in the PI sample but the signal is weaker. Gel A was run with 15ug protein, 50ng CotI and 0.01 units poly(dI-dC). Gel B samples contained 30ug protein, 50ng CotI and 0.03 units poly(dI-dC).

Following optimisation of sample conditions, the EMSA was performed as shown in figures 6.15 and 6.16. There is a clear absence of protein/probe complex 2 from both the non transfected cell nuclear extract (NT) and from the empty expression vector extract (0). This indicates that protein/probe complex 2 is indeed the FOXF2 peptide. Further evidence for the identity of this band can be seen in figure 6.16 in which the band produced by nuclear extract taken from cells transfected with the

polyglycine expansion mutant FOXF2 (PG) is clearly emigrating slower than either the wildtype (WT) or V410F mutant (VF) forms of the peptide.

The key finding from this analysis however, was the absence of a clear band in the sample that contained nuclear extract of cells transfected with the FOXF2^{W174R} mutant peptide. This suggested that DNA binding by the protein with this mutation is either very weak or absent and could explain the primary cause of the deficiency in gene function that leads to the phenotype of mutant individuals observed in chapter 5.

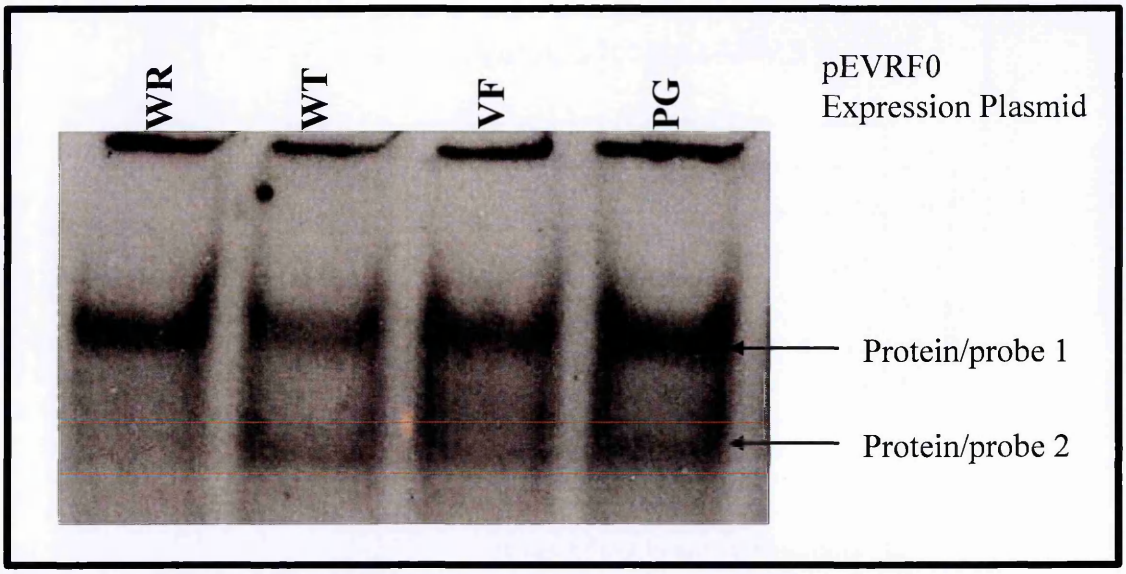


Figure 6.16: Enlarged view of EMSA
Variable band intensities from nuclear protein extracts taken from cells transfected with different isoforms of a FOXF2 expression plasmid. The red lines are positioned to emphasise a subtle elevation of the protein/probe 2 band in the polyglycine mutant sample ($FOXF2^{PG}$). There is no clear band in the nuclear extract taken from cells transfected with the FOXF2W174R plasmid, indicating a reduced affinity for the probe sequence.

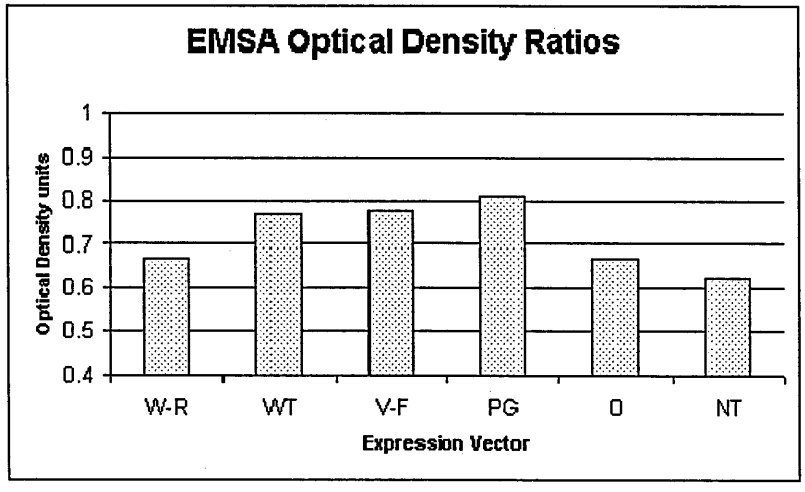


Figure 6.17 Optical Densities of EMSA bands
The optical densities of the bands hypothesised to result from FOXF2-probe complexes in figure 6.13 A. The background intensity of the membrane was 0.44. A lower signal strength in the W174R sample compared with the other FOXF2 construct transfections can be seen although there is a high background signal in each lane that hinders this analysis. The W174R sample showed an optical density of 0.6-0.7 OD which is the same range as the empty expression vector (0) and the non-transfected control (NT). This is likely to be caused by non-specific binding of the DIG labelled probe to other constituents of the samples.

In figures 6.15 & 6.16, an upper potential protein/probe band is seen in each of the samples including the non transfected and empty vector transfected extracts indicating the possible presence of an endogenous forkhead protein in the Cos7 nuclear extracts. The lower band (protein/probe 2) is present in the wildtype, V410F mutant and polyglycine mutant extracts but was not detected in the empty vector or non-transfected extracts. The W174R mutant sample does not show a clear second band as seen in the other FOXF2 transfected samples although a faint smear can be seen which stops at the same level as the second band in the other samples.

The absence of the second band in the empty expression vector and non-transfected control nuclear lysates; and the bands strong presence in all FOXF2 transfected

cultures except for the FOXF2^{W174R} isoform, provide credible evidence that this band was FOXF2.

Antibody assays though, did not provide such consistent evidence. Additional bands were seen in the antibody containing samples in figures 6.14-6.15, but no bands could be seen during an earlier attempt to optimise the signal strength generated by adjusting the total protein concentration within the antibody/nuclear extract mixture (figure 6.12).

One difference between these mixtures was the absence of DNA loading dye in the supershift optimisation gel (figure 6.12). It is therefore possible that the DNA was less likely to sink to the bottom of the wells when loading the samples without loading buffer. However the reason that loading dye was initially excluded from the binding mixture was because the glycerol in the binding buffer was considered sufficient to bring the sample to the bottom of the well during loading and it was considered preferable to remove as many potentially interfering factors from the binding mixture as possible before experimental runs. Despite the lack of colour in either the sample or the probe, the differing viscosities of the binding buffer and the surrounding running buffer allowed visualisation of sample loading albeit with less clarity; all samples appeared to be correctly loaded into the wells.

The banding pattern in the preliminary supershift optimisation (figure 6.12) matched samples without antibody. Subsequent samples which contained loading dye, did

exhibit additional bands on addition of the antibody (figure 6.14). It is difficult to speculate why the addition of loading dye may have affected the run profile. It seems unlikely that the recommended EMSA loading dye would be inconsistent with antibody binding and even if this were the case it would be a remarkable coincidence if the loading dye corrected the alteration that had previously caused the loss of additional bands.

It is possible that the additional bands seen with the supershifts weren't complexes of antibody and FOXF2. There was a weak banding pattern which reflected the band sizes seen in the apparent positive samples when antibody was added to the DIG probe alone. This indicated that non-specific binding of the antibody may have occurred. Although the signal is weaker in this control lane than in the other antibody containing lanes, this feature is not necessarily reassurance that the binding is not non-specific. Nuclear extracts would contain higher concentrations of material which could act as a substrate for non-specific binding.

In gel 6.15B of the EMSA, the upper band is present in all samples but is strongest in the empty vector and non-transfected samples, where there are no second bands. In the wildtype, V410F mutant and polyglycine mutant samples, the second bands are stronger than the top bands. There does not appear to be a second band present in the W174R mutant sample. The second band was very slightly larger in the polyglycine mutant sample (figure 6.16), providing further evidence of its identity as the product of the transfected vector.

The relatively weaker top bands in those samples in gel B that also contain a lower band suggest that there may be an inverse relationship between the strength of the top band with the strength of the lower band. One possible explanation for this is that when FOXF2 is absent or is present in a form with reduced sequence affinity, the probe may be bound more efficiently by this other protein. If this is the case, when FOXF2 is present, it could provide direct competition for the available labelled sequence and thus reduce the signal from the other protein. However the probe is present in excess in the binding mixture judging by the strength of signal from free probe in all samples. This would suggest that the effects of differential affinity should be minimised.

The lack of a visible, distinct lower band in the W174R sample in either set of binding conditions suggests that binding efficiency is dramatically reduced, and possibly completely disrupted in this mutant form of the protein. Although the ratios for the W174R mutant and the empty expression vector show significantly higher band intensities than the image background, analysis of the image shows a smear between the top band and the position at which the lower potential FOXF2 band is found in those samples which have a distinct band at this position. This smear is present in all samples whether a distinct lower band is visible or not.

The smear could result from degradation of the protein represented in these images by the top band, or it may result from other proteins which demonstrate some level

of specificity to the binding sequence, or proteins with a non specific affinity to this sequence. Whatever the cause of this smear, its presence in the band shift suggests that it is not indicative of FOXF2 binding. For the purposes of this investigation, this signal can therefore be regarded as background noise.

Previous structural predictions of forkhead proteins suggest that one of the α-helices which are common to all of these proteins, fits into the DNA major groove and one of the two wings has contact with the minor groove at the 5' end of the core sequence [95]. This suggests a requirement for precision in the DNA interaction of forkhead transcription factors and it is possible that an alteration to the structural integrity of the molecule could disrupt the arrangement necessary for DNA binding.

The FOXF2^{W174R} mutation has occurred at the centre of one of the structurally significant β-sheets (See figure 4.17). The wildtype tryptophan residue is not predicted to have direct contact with DNA during the interaction of FOXF2 with its target sequence but because the morphology of the protein could be drastically altered by this mutation – the FOXF2^{W174R} isoform may not function normally.

6.3 Transactivation Assay

Transactivation Assays are useful tools for examining the interaction of transcription factors with specific DNA sequences and the transcriptional machinery. The cDNA of a transcription factor of interest is incorporated into an expression vector and the target sequence is attached to a gene directing the synthesis of a reporter molecule. Reporters are required to be easily detectable and quantifiable molecules such as luciferase or alkaline phosphatase [275]. Levels of this molecule can therefore be used to ascertain the activational behaviour of the gene product of interest on DNA target sequence.

This test was considered appropriate for preliminary analysis of the molecular effects of the V412F mutation and to contribute towards the understanding of how and why the W174R mutation has an effect on the phenotype of the animal. It was hoped that it may clarify what change occurs in the activity of the W174R FOXF2 protein and could also determine whether an alteration in the transcriptional activity of the protein occurs in mutants

6.3.1 Construction of the Reporter Plasmid

The pTAL-SEAP reporter vector contains a modified alkaline phosphatase gene, which encodes a secreted form of alkaline phosphatase (SEAP). The promoter (pTAL) is a region from the Herpes simplex virus thymidine kinase (HSV-TK) promoter. The secreted alkaline phosphatase can be taken directly from the cell media for quantitative chemiluminescent assays of gene activation. This system

therefore has the advantage that cell lysates are not required and several time points can be taken for a single batch of transfected cells.

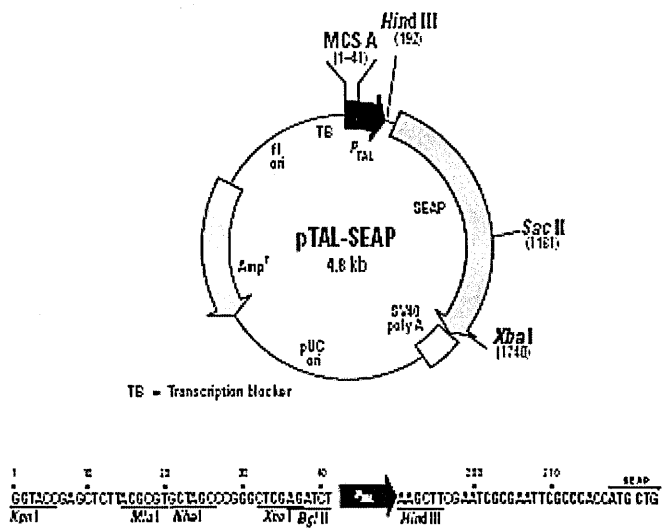


Figure 6.18: pTAL-SEAP Reporter
The pTAL-SEAP vector and restriction sites available for the insertion of binding regions.

The restriction sites MluI and BglII were selected from the multiple cloning site (MCS) as appropriate insertion loci, due to their identical optimal activation conditions (see section 2.2.12) and also because they are not immediately adjacent to each other within the MCS. Both of these factors facilitated their use in a double digest.

Two Primers were adapted from those used by Miura et al [241], incorporating the FOXF2 recognition sequence which had previously been determined by Pierrou et al [205], but replacing the HindIII and SalI restriction sites with MluI and BglII sites (see appendix 6.1.5 and figure 6.21).

These primers that contained the FOXF2 recognition sequence were intended for multimerisation to produce a reporter insert containing a multiple copy binding site as illustrated in figure 6.20.

However, although multimerisation of the oligonucleotide was induced as can be seen in figure 6.19, the multimerised product was not successfully incorporated into pTAL-SEAP. Colony PCR with the pTAL-SEAP primers (appendix 6.1.1) of 130 clones that resulted from transformation of XL1-Blue cells with the ligation product of the multimer and linearised plasmid was carried out following Qiaquick gel purification of the plasmid DNA. Ligations had used the X5, X7 and X9 product sizes, and T4 ligase concentrations of 1.6u/ul and 4u/ul. All of the PCRs resulted in a 340bp product which indicated that incorporation of the multimer into the pTAL-SEAP plasmid had not occurred in any of these clones.

One likely explanation for the failure to identify a recombinant pTAL-SEAP plasmid containing the multimer is that the purification method may not have been 100% efficient in the removal of restriction enzymes. This would have resulted in the digestion of the multimer during the ligation step.

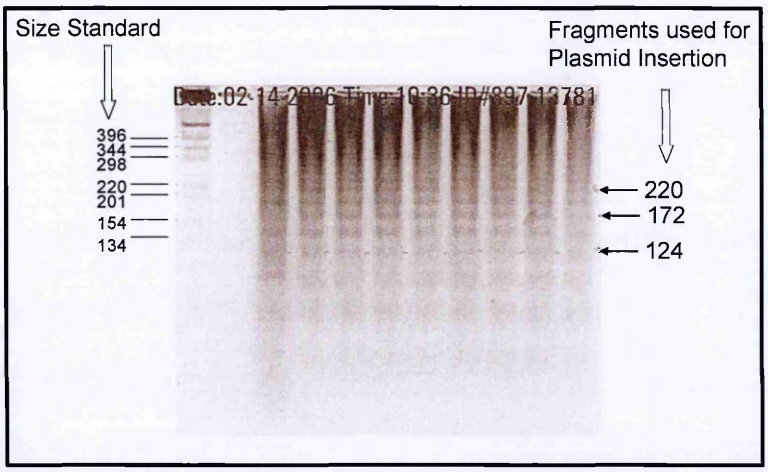


Figure 6.19: Multimerisation of the single binding site oligonucleotides
Band sizes suggesting they contained multimers with odd numbers of binding sites (124bp= X 5 copies; 172bp= X 7 copies;220bp= X 9 copies) were selected because BglII and MluI terminal restriction sites were required at either end of the multimer for incorporation into the plasmid. These bands were cut and gel purified as described in section 2.2.8. This product was then added to linearised pTAL-SEAP in a ligation mix as described in section 2.2.17. None of the plasmids that were analysed for size differences on agarose gels indicated the insertion of any of these products.

Following the difficulty in inserting the binding site multimerisation product, another pair of complementary oligonucleotides was designed which included five FOXF2 binding sites. A Mlu I site at the 5' end of the forward sequence, and a Bgl II site at the 5' end of the reverse sequence were incorporated into the design of these oligonucleotides.

A "spacer" sequence of TTC (GAA on the reverse strand) was also added between each binding sequence on the forward sequence to add space between binding sites and reduce the possibility of an inhibitory effect by binding at one site to binding at adjacent sites (These oligonucleotides are identified as FOXF2_BS in appendix 6.1.5 and are illustrated in figure 6.21). However, it should be noted that the space

between binding sites on these oligonucleotides (3 bases in length) is less than the space that occurs between binding sites on the multimer (figure 6.21).

Of the 68 colonies that were produced from the transfections which were performed in the construction of the pTAL-SEAP reporter vector with FOXF2 binding sites, 4 colonies gave products in the range of the 441 bp that would be expected if the binding site oligonucleotide was inserted into a pTAL-SEAP cloning site. Three of these were from the 300:1 oligonucleotide to plasmid ratio and 1 of them was from the 30:1 oligonucleotide to plasmid ratio. On sequencing this product, all of these clones contained the binding site insert and one of them was selected for transfection of Cos7 cells (as described in sections 2.2.33-2.2.35).

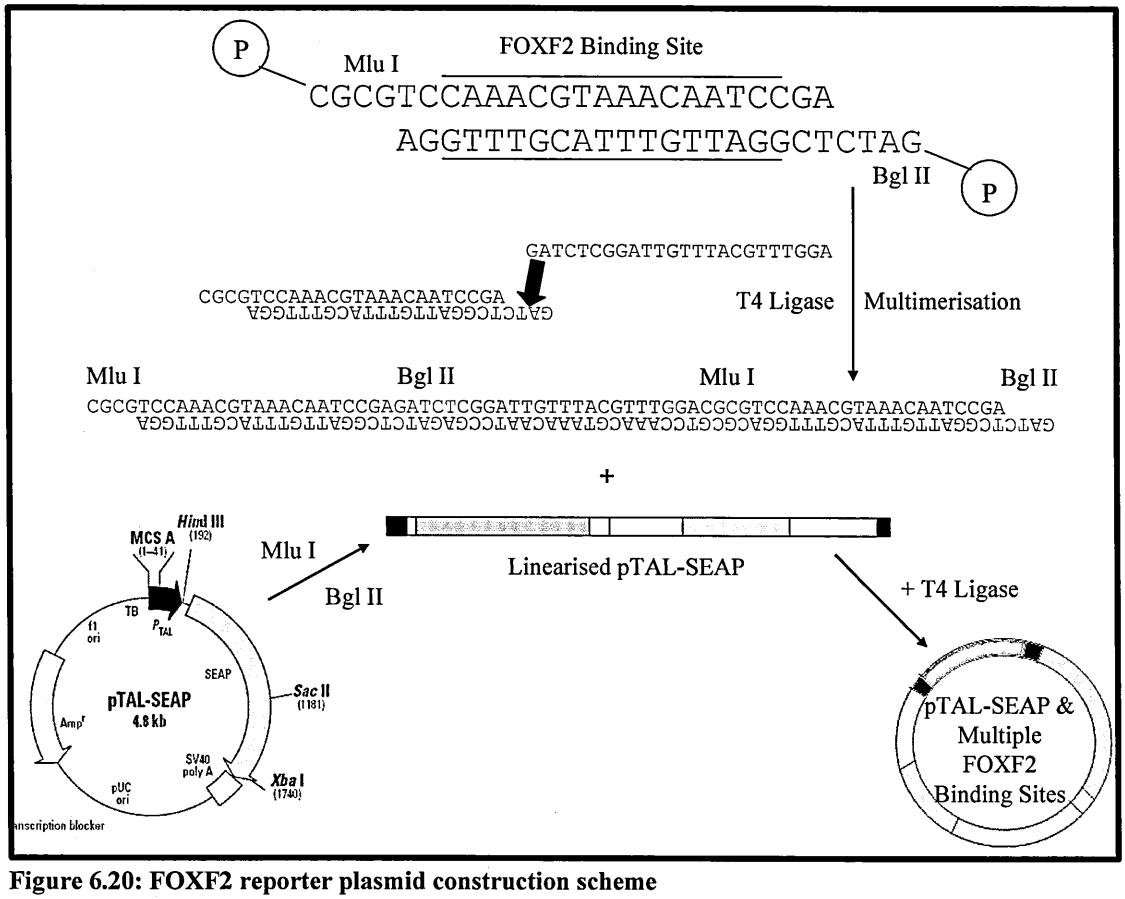


Figure 6.20: FOXF2 reporter plasmid construction scheme
Scheme illustrating the first method for the construction of a FOXF2 reporter plasmid. Phosphate labelled oligonucleotides which contain the FOXF2 binding site and either a Bgl II or Mlu I restriction site are multimerised with T4 ligase to form double stranded DNA containing multiple FOXF2 binding sites. The pTAL-SEAP plasmid is linearised with Bgl II and Mlu I before purification with a Qiaquick DNA purification kit. The multimer is then ligated to the linearised plasmid with T4 ligase.

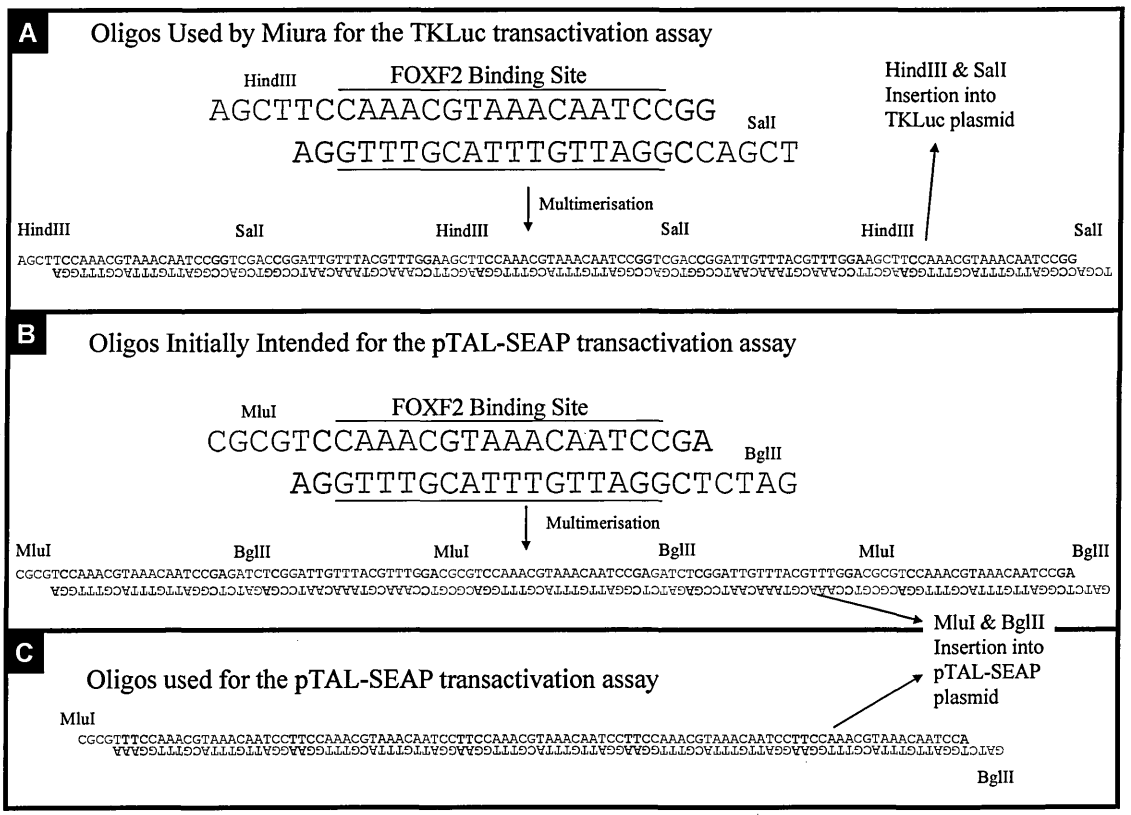


Figure 6.21: The different oligonucleotides used for binding site insertion
A comparison of the oligonucleotides and their products for 3 FOXF2 transactivation assay schemes. Blue text is FOXF2 binding sequence. Red text is restriction sequence. A: The previously reported oligonucleotides and 5X multimerisation product incorporated by Miura into the TKLuc reporter vector [241]. B: The oligonucleotides were adapted to allow insertion into pTAL-SEAP by replacing the HindIII and SalI sequence at the 5' ends, with MluI and BglII respectively. A guanine at the 3' end of the MluI oligonucleotide was also changed to an adenine to allow annealing to a thymine in the BglII restriction sequence. Nucleotides that vary between these sequences and those designed by Miura are highlighted in yellow. C: Extended oligonucleotides containing 5X FOXF2 binding sites that eliminated the requirement for multimerisation and therefore also eradicated the problems associated with the insertion of a multimerised product. The space between binding sites in this double stranded product is only 3 nucleotides in length in contrast to the 8 nucleotides between binding sites in the two multimerisation products.

The transactivation assay was intended to highlight whether there was a difference between the activational activity of the different FOXF2 isoforms. In this study, SEAP production is the quantifiable molecule that provides a relative measure of the activation profiles of the different FOXF2 isoforms. SEAP levels were significantly higher when the reporter plasmid was co-transfected with the wildtype FOXF2 expression plasmid than with the W174R mutant form (figure 6.23 & appendix table

6.10). However this was the case with both the empty pTAL-SEAP vector and with the reporter containing the binding site. Levels of SEAP were lower in transfections with the FOXf2 binding site reporter than in transfections of the empty vector (figure 6.23). It therefore appears that the pTAL-SEAP vector may be activated by FOXF2 in the absence of the binding site.

Large differences between SEAP samples were seen in the EMSA control data, causing early EMSA results to be discounted - the lack of signal in the W174R signal correlated with low SEAP levels. Differences in SEAP activity were several orders of magnitude greater in the EMSA flasks than in the smaller wells used for transactivation assays with the same vectors. The unexpectedly high levels of these empty SEAP vector results could be considered as further evidence that the empty reporter vector may be activated by FOXF2 (see figure 6.22 and table 6.7).

SEAP levels in the media of $FOXF2^{V410F}$ transfected cells were higher than in the media of $FOXF2^{WT}$ transfected cells when co-cultured with PTAL-SEAP. This was in contrast to the transactivation assay which indicated a slight reduction in levels of $FOXF2^{V410F}$ SEAP activation when compared with wildtype levels.

Amounts of the transfection reagents were considerably larger for the EMSA flasks than for the wells of the transactivation assay due to the larger vessel sizes. However none of the conditions or reagents were exactly proportional (appendix table 6.11) and this will have affected the activity of the FOXF2 protein.

Expression vector volumes in the EMSA transfections were added at equivalent levels. pTAL-SEAP volumes added to the transfection reaction were ~4X greater in the EMSA, but ~5X greater in the transactivation assay when calculations were corrected for surface area.

These volumes were different because of the transfection conditions for the different vessel types that were recommended by Merck Biosciences, but were also due to different requirements of the control vectors in the two experiments. For example, in the EMSA, pTAL-SEAP acted as the transfection control whereas for the transactivation assay – GFP filled this role because pTAL-SEAP had a more central role in the experimental setup.

One important point to consider is that the experimental design for the EMSA transfections did not incorporate pTAL-SEAP to obtain stand-alone experimental data but as a transfection control. The inconsistent SEAP production and possible activation of pTAL-SEAP by the FOXF2 peptide complicates the trans-activation picture, but has allowed an additional dimension to the data analysis.

6.3.2 Control Levels in EMSA Transfections

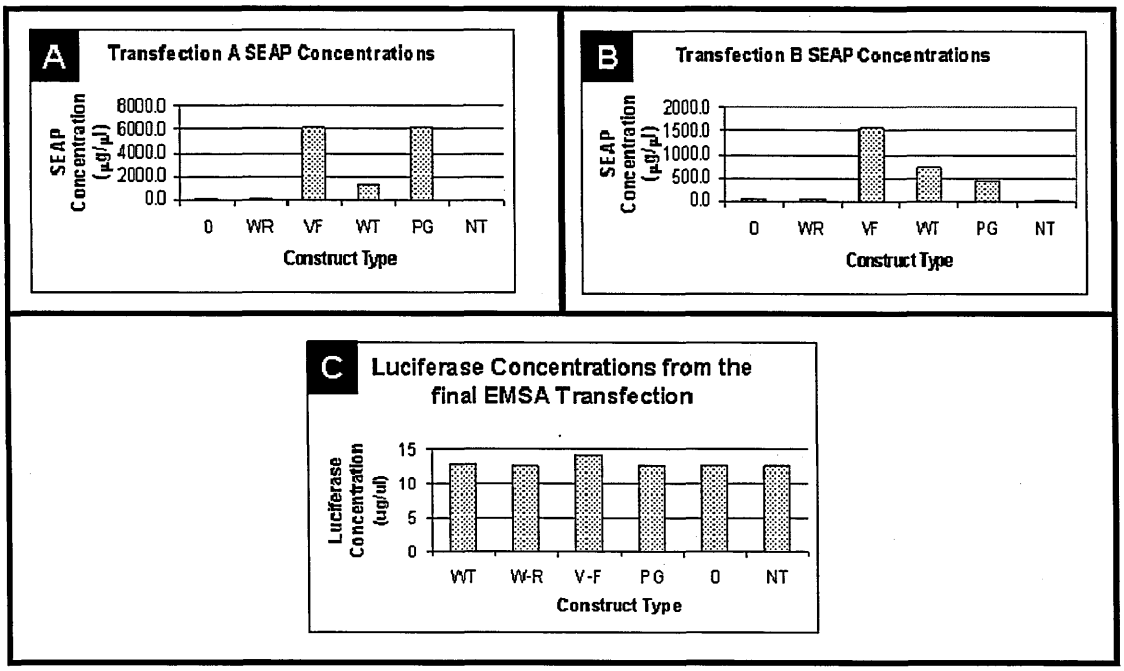


Figure 6.22 Transfection efficiency control data

Figure 6.22: Transfection efficiency control data. SEAP data from two different EMSA transfections in which pTAL-SEAP was used as a transfection control and one in which pMLuc-3 CMV was the control. The large variability between the samples in A led to a different approach used for the transfection of the samples in B: In A, the genejuice/vector complex was added to the cells directly, by tilting the T175 flask and pipetting the complex onto the cells. However, complete coverage of the cells with the complex was not achieved in any of the cultures before recovering the cells with media and mixing, and there were therefore concerns that this uneven distribution of the genejuice/vector complex may have affected the transfection efficiency within each sample. The transfection reagents in B were prepared in the same way, but they were added to the media at the bottom of vertical flasks before bringing them to a horizontal position then mixing by tilting from side to side. This approach was also applied to the transfection shown in C, in which pMLuc3-CMV was used as the transfection control. The variability between sample controls was greatly reduced in this transfection – meaning that the EMSA gels could be analysed without concerns about data variability and reliability. The EMSA gels shown in figures 6.15 & 6.16 utilised the pMLuc3-CMV transfection control.

6.3.3 SEAP Data

SEAP levels increased in the presence of $FOXF2^{WT}$ and $FOXF2^{V410F}$ expression plasmids above the levels associated with the empty expression vector, but the only significant increase when co-transfected with the binding site containing reporter, according to the Tukey test was the wildtype FOXF2 transfection at 24 days.

There were other more significant increases when the binding sites were absent from the reporter, when empty pTAL-SEAP.

Transcriptional activity was increased compared to the empty reporter vector in every sample in which the FOXF2 binding site was present in pTAL-SEAP. WT shows the strongest increase, followed by V410F. The W174R mutant doesn't demonstrate an increase in transcription levels in the reporter containing the binding site, but there is an increase much higher than those seen with the empty expression vector.

A two-factor ANOVA with replication was performed on normalised SEAP data for each of the three time points. These data were analysed separately in order to maximise the probability of detecting a SEAP production differentiation between samples.

There were highly significant differences within both the reporter ($6.28 X10^{-6}$ to $3.02 X10^{-5}$) and expression plasmid groups ($1.02 X10^{-8}$ to $4.37 X10^{-6}$). In both cases the most significant differences occurred at 24hrs.

This test does not however show where the differences occurred, and it was therefore necessary to perform subsequent post-hoc tests. A Tukey test was used to compare each transfection with the other transfections (appendix table 6.10).

Previous studies have shown that pTAL-SEAP without modification has a very high basal promoter activity at ~6650 RLU/10^5 cells [276]. It could be feasible that the levels of SEAP with the empty pTAL-SEAP vector were at background levels. However, this hypothesis would ignore the significant difference of SEAP production between the different expression plasmids.

The Tukey test demonstrated that there was a highly significant difference ($p<0.01$) between the empty pTAL-SEAP and the binding site reporter when co-transfected with the wildtype FOXF2 expression plasmid. The difference between the reporters was less marked in the presence of the V410F isoform, with significance levels around the $p=0.05$ level ($q\sim4.9$) at each time point and were marginally significant ($p<0.05$) at 24 and 48 hours and marginally insignificant ($p>0.05$) at 72 hours.

The empty and binding site reporter vectors however do not show a significant difference in SEAP production when co-transfected with either the empty expression vector or the FOXF2^{W174R} isoform.

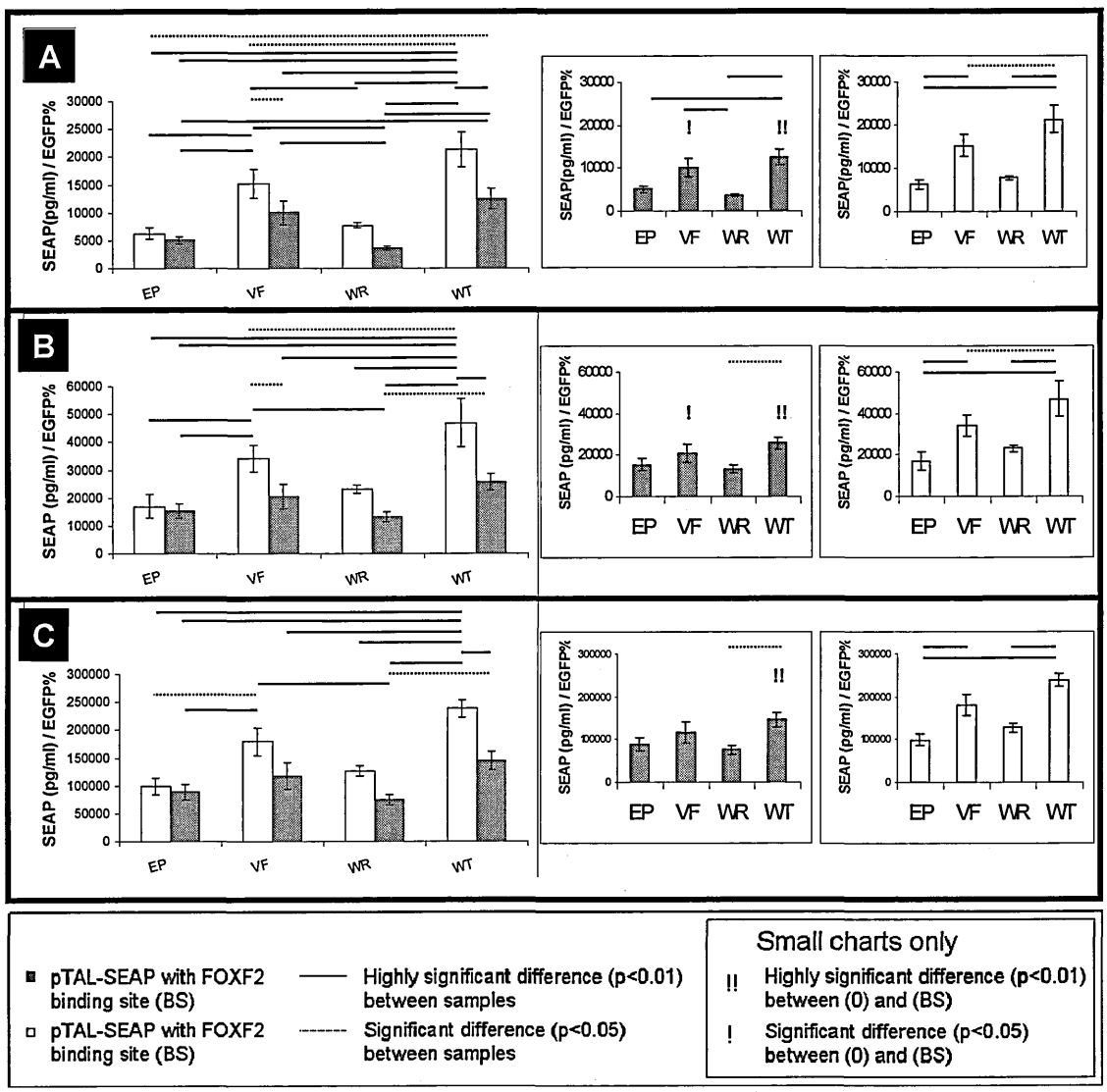


Figure 6.23 Transactivation assay data
Normalised SEAP Levels for each expression vector in the presence of either empty pTAL-SEAP or pTAL-SEAP with the FOXF2 binding site insert. (A) normalised SEAP levels at 24 hours, (B) at 48 hours and (C) at 72 hours after transfection. The larger charts represent each of the significant differences between samples from either reporter plasmid whereas the small charts represent significant differences between expression plasmids transfected with the same reporter. Error bars represent Standard deviation from three cultures per sample. The horizontal lines indicate significant differences between the samples at either end as determined by Tukey test. !=significant difference between reporters for the same expression vector ($p<0.05$), !!=highly significant difference between reporters for the same expression vector ($p<0.05$).

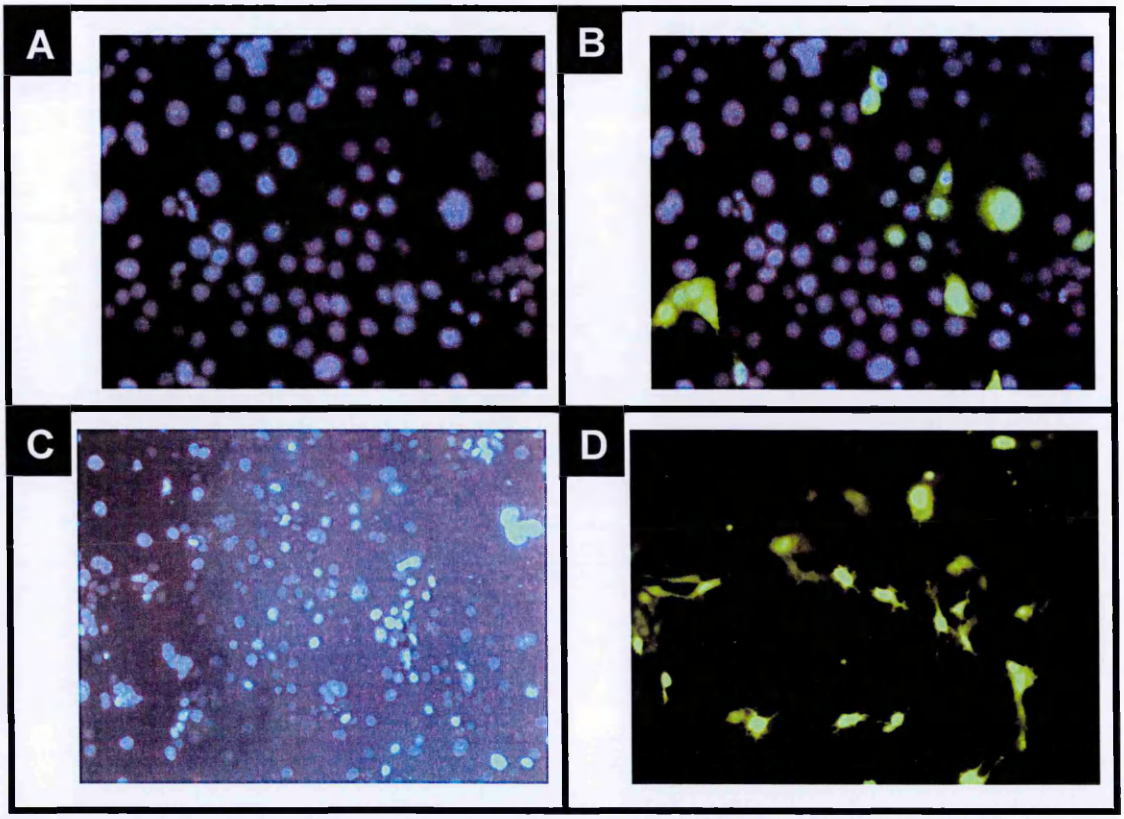


Figure 6.24: EGFP was used a a transfection control in the transactivation assay
Views with a blue filter (A + C) which highlight DAPI stained cell nuclei alongside identical view fields with a green filter (B + D) which allows the visualisation of cells transfected with EGFP. Transfection ratios were determined by comparing the numbers of GFP transfected nuclei with the blue DAPI stained nuclei. This ratio was then be used to assess the transfection efficiency for each culture thereby allowing the normalisation of reporter yields. Images A+B were taken from an Olympus BX40 Microscope. Green and blue wavelengths were filtered electronically with Smartcapture (Digitalscientific UK Ltd) following image capture. The EGFP view could then be superimposed onto the DAPI view as in B. Photographs C + D were obtained with a Zeiss Axioplan 2 Microscope using green and blue physical filters. These images were processed with the software Axiovision v4.5.

6.4 Potential pTAL-SEAP Endogenous FOXF2 Binding Sites

One possible explanation for the upregulation of alkaline phosphatase in those transfections which contained empty pTAL-SEAP is that this plasmid already contains a number of potential binding sites for wildtype FOXF2.

The core consensus Forkhead Binding Sequence that was determined by Pieirou et al [205] is: RTAAAYA

R = Purine = Adenine & Guanine (A + G)

Y = Pyrimidine = Thymine & Cytosine (T + C)

A

Consensus Core sequence:	**Forward**	**Reverse**
	RTAAAYA	TRAAATY
	ATAAATA	TATTTAT
	GTAAATA	TATTTAC
	ATAAACA <	TGTTTAT <
	GTAAACA	TGTTTAC

B

1. gcaATAAACAa (pTAL-SEAP 1 at position 1886-1896 within the vector sequence).
2. aaaATAAACAa (pTAL-SEAP 2 at position 4018-4028 within the vector sequence).

3. aacGTAAACAa (The optimum binding sequence determined by Pierrou)

a t

C

	Nucleotide Position Frequencies (%) of Random Oligonucleotides that bound to FOXF2 Protein										
	-3	-2	-1	1	2	3	4	5	6	7	+1
G	7	18	30	74							7
A	63	48	11	26		100	100	100		100	59
T	15	15	18		100				7		18
C	15	18	41						93		15

Figure 6.25: Potential Foxf2 binding sites on endogenous pTAL-SEAP

Figure 6.25: Core consensus binding sequence recognised by Fox genes, determined by Pierrou et al [205] - and their similarity to two potential FOXF2 binding sites within the pTAL-SEAP cloning vector. The Binding site in each figure is highlighted in yellow. A= The Forward and complementary sequence of the consensus forkhead binding sequence. Taking into account the different possible nucleotides at the purine and pyrimidine positions, the consensus binding sequence could take the form of any of the sequence combinations listed below the core sequence. The arrow points towards the sequence that was identified in pTAL-SEAP. B= The binding sequence determined by Pierrou for Foxf2 at B3 is compared with the two potential forkhead binding sites from the pTAL-SEAP sequence at B1 and B2. pTAL-SEAP potential binding sequence 1 is at nucleotide position 1886-1896 within the listed vector sequence (http://www.clontech.com/images/pt/PT3271-SEQ.pdf for sequence and http://www.clontech.com/images/pt/PT3271-5.pdf for feature map). This region contains the SV40 late mRNA polyadenylation signal at position 1888–1893. pTAL-SEAP potential FOXF2 binding sequence 2 is at position 4018-4028 of the vector. This is 3' of an ampicillin resistance promoter which lies at 4012-4007. C=Data extracted from Pierrou et al [205] showing the frequency of nucleotides present at positions in relation to the core sequence (yellow highlight) within 27 random oligonucleotides that bound to FOXF2 protein.

One of the four potential sequence combinations can be found within the pTAL-SEAP Vector Sequence at two separate locations as shown in figure 6.25B.

In the FOXF2 binding site study discussed in section 6.1.2, none of the nucleotides identified by Pierrou were exclusively present at any position in the flanking sequence. This is in contrast to the nucleotides within the core sequence where only oligonucleotides containing the listed sequence bound successfully to the protein (see figure 6.25C). This indicates that the flanking sequence tolerates a greater degree of variation than the core sequence.

Both of the pTAL-SEAP potential binding sequences contain a sub-optimal nucleotide at position 1 of the core sequence. As can be seen in figure 6.25C, this Purine specific position was occupied by a guanine in 74% of Pierrou's binding oligonucleotides whereas both potential sites on pTAL-SEAP contain an adenine at this position as was present in 26% of the oligonucleotides.

The two pTAL-SEAP sequences contain an identical FOXF2 binding sequence at their assumed core although their flanking sequence differs. As can be seen in figure 6.25B, pTAL-SEAP potential binding site 1 contains the sequence GCA at position -3 to -1; binding site 2 contains three adenosines at these positions relative to its core sequence. The optimal flanking sequence at this position according to Pierrou was AAC, which only differs from the 2nd pTAL-SEAP sequence by one base – in which the nucleotide immediately 5' of the core sequence is an Adenosine

In the oligonucleotide binding experiments – An A was present at this position in 11% of the oligos that bound to FOXF2; lower than the percentage of any other base. C was seen in 41% of the oligos at this position, G was in 30% and T was in 18%.

6.5 Conclusions

The high levels of SEAP that were associated with the empty pTAL-SEAP vector may be due to the presence of potential endogenous FOXF2 binding sites. Two such sites were identified on analysis of the pTAL-SEAP sequence. This explanation is given further credibility by the enhanced SEAP activation when empty pTAL-SEAP was co-transfected with a wildtype FOXF2 expression vector. The conditions in this experiment resulted in a considerably larger difference between wildtype FOXF2 and control vector when co-transfected with empty pTAL-SEAP, providing more evidence that FOXF2 is activating the empty reporter.

Promoter sequence of Prostate specific antigen (PSA) has been shown to increase in activity when duplicated on a luciferase reporter but this activity was reduced with the addition of a third copy [277]. If the problem with the FOXF2 binding site reporter is not specific to the background vector but rather to the insert sequence, it is possible that binding site copy number or perhaps the proximity between binding sites may inhibit the activity of the transcription factor to below the background activity that may be encouraged by endogenous FOXF2 binding sites.

The W174R mutation appears to severely reduce the affinity of FOXF2 to its recognition sequence. Although antibody binding to FOXF2 was not achieved with the Western blot experiments, only one of the two available antibodies were tested with this approach and this antibody also failed to produce an alteration to the suspected FOXF2 band in the EMSA when added to test for a supershift. The other antibody that showed some potential in the supershift experiment (FOXF2 antibody 1) was not applied to a western blot of the nuclear lysate.

6.5.1 Potential Sources of Error that could Explain the Data

Concentration calculation errors are unlikely as the three repetitions were performed simultaneously, so variability of sample and reagent concentrations is likely to be minimal. One possible explanation for the erroneous data that is more likely given the simultaneous sample preparation, is that the two reporters were mixed up before their addition. However, high levels of activation were seen in two preliminary EMSAs that used SEAP data as transfection controls. This evidence in support of the endogenous pTAL-SEAP binding sites also goes some way to preclude the possibility that reporters were incorrectly added.

SEAP levels detected in the EMSA transfection controls were ~2 X 10^8 higher than those of the transactivation assay. The total amount of pTAL-SEAP added to each EMSA transfection was increased by ~4 fold when compared to the amounts added to the transactivation assays, but reduced by a factor of 5 when the volumes were

corrected for surface area of the respective culture vessels. The higher SEAP levels in the EMSA suggest that the transfection reagents were added at sub-optimal levels in the transactivation assay.

Preliminary optimisation of the transactivation assay was performed prior to the investigation proper – but this involved measurements of GFP vector transfection ratios with different concentrations of genejuice rather than specific optimisation of pTAL-SEAP. It appears likely that optimal conditions for the transfection of pTAL-SEAP are different from those of GFP.

It is interesting to speculate that greater significance levels may have been obtained if transfection efficiencies were higher, particularly for the V410F plasmid for which in the EMSA there was such a large difference compared to SEAP production from the empty expression plasmid.

The fact remains though that even in the EMSA, where pTAL-SEAP transfection conditions appear to have been closer to optimal, SEAP production by the empty pTAL-SEAP vector is higher in the presence of wildtype FOXF2 than is pTAL-SEAP with multiple copies of the FOXF2 binding site. If this is genuine activation of the empty reporter, pTAL-SEAP should be excluded from further investigations into FOXF2 activity unless the interaction itself was of interest.

6.5.2 Future Work

The lack of variation in the luciferase production of cells transfected with pMLuc-3 CMV (figure 6.22 C) suggests that this reporter may not contain endogenous FOXF2 binding sites . One experiment that could be used to further investigate the elevated SEAP levels in the empty pTAL-SEAP transfected cultures would be the introduction of the putative pTAL-SEAP binding sites into pMLuc-3 CMV. This could be compared with transactivation data from a pMLuc-3 CMV vector containing the original binding site insert to see if the improvement in transfection efficiency with the pTAL-SEAP endogenous sites is genuine.

The difference between SEAP data for the V410F transfections in the EMSA and in the trans-activation assay suggests that the higher initial concentrations of pTAL-SEAP in the transactivation assay were sub-optimal. Given the large SEAP increase in the EMSA samples, it could be useful to repeat the transactivation assay at the same relative concentrations as the EMSA to see if those results are repeated.

In the conditions used for obtaining protein for the EMSA, this mutation appears to increase the transcriptional activity by ~2-5 fold when in the presence of the empty reporter. Although the SEAP data from these two experiments cannot be directly compared with the transactivation data due to the difference of scale, the relative difference in activity compared to other expression vectors is of potential interest and further analysis in different cell concentrations could help to further characterise this mutation.

These potential future experiments are only worthwhile if the FOXF2 activation of unaltered pTAL-SEAP was genuine. The construction of a pMLuc-3 CMV reporter containing binding sites is necessary for the investigation of the pTAL-SEAP endogenous binding, but a more direct approach to fulfil the initial aims would be to repeat the assay with this vector. An investigation of the pTAL-SEAP binding sequences may be worthwhile if as would appear from the preliminary SEAP data that was obtained, they are more efficient activators of the protein than those sequences that were determined by Pierrou et al.

Chapter 7: Discussion

An accurate physical map of the Del 36 H region has facilitated the mapping of mutations arising from the regional screen with a greater resolution than would have otherwise been possible. It has also allowed the precise identification of the deletion breakpoints, thereby clarifying which of the peripheral genes are deleted.

The minimal tiling path that was selected from the physical BAC map made a direct contribution towards the mouse genome sequencing effort and therefore the Ensembl sequence (http://www.ensembl.org/Mus_musculus/). Del36H was one of the first regions that was sequenced using the hierarchical rather than the whole-genome approach and the sequence quality will have therefore benefited from the improved confidence in the sequence integrity of the large repetitive regions within the interval [43].

A search for polymorphic markers had the fortunate by-product of large amounts of haplotype map data for the Del36H region. At smaller than 668kb, the average haplotype block size for this region was appreciably smaller than the genome overage of ~3.57Mb that Wiltshire [226] has previously described. Wiltshire also noted that the borders of haplotype blocks are associated with recombination hotspots. The smaller blocks within Del36H would therefore suggest that recombination is likely to occur at high frequencies within the region and this may explain the larger size that the genetic map predicted for the region.

One key benefit of the intensive characterization of this region was the potential for an improved understanding of the syntenic region on human chromosome 6p that is deleted in 6p deletion syndromes. Patients with these chromosomal anomalies exhibit a wide range of physiological and neurological phenotypes; the variation of which is caused primarily by the variability in size and positioning of the deletions.

There has been an effort to correlate the varying extents of these deletions with the associated phenotypes [2, 75, 80]. This will have implications beyond these rare deletion syndromes – there are also many non-syndromic disease loci within the region and an analysis of different monogenic mutations in mouse models should aid in the identification of the associated causative genes.

The absence of coding mutations in Foxf2 for any of the lineages derived from the lethal recessive screen implies that there are no mutant FOXF2 proteins causing a phenotype, but it does not rule out mutations that effect expression of the gene. There are regions of both the 3' and the 5' untranslated regions and promoters that were not sequenced, although they may be important for the gene's correct regulation of expression.

In-situ hybridisations of Foxf2 did not highlight any changes in the expression of Foxf2 in any of the mutants [26]. However this does not rule out the possibility that there may be subtle expression alterations.

7.1 Physiological Phenotyping

One of the most significant findings of this project was that a mutation in Foxf2 can cause lethality without inducing a cleft palate. Previous characterisations of Foxf2 have involved the analysis of the Foxf2 knockout [153, 154]. Wang observed no physiological difference between the mutant and wildtype that could explain the deaths of the knockouts other than a deformed tongue and cleft palate which was suggested as a possible cause of the associated gas-distended GI tract.

It is possible that these characteristics are the immediate cause of death in the knockout animals – but there was no evidence for any of these features in the $Foxf2^{W174R}$ mice. This opens the door to an expanded investigation into the cause of $Foxf2^{W174R}$ mortality and will allow the characterisation of other phenotypes which are impossible to identify in the knockouts due to the overwhelming incidental effects of the cleft palate (e.g.: the gas-distended gut). This mutant therefore provides an immediate vindication of the use of the ENU archive gene-driven-screen approach for the identification of novel functional mutations.

It should be emphasised that the genetic backgrounds of the Foxf2 knockouts and Foxf1/Foxf2 compound heterozygote knockouts used by other laboratories were of C57BL/6 origin [153, 154]. This is in contrast to the majority of $Foxf2^{W174R}$ phenotyping which was performed on mice that had a majority C3H background but with a significant minority (0.4%-6.3%) of BALBc depending on the generational gap between founder male and test-subject (lineages were founded by mutagenised

BALB/c mice that were then outcrossed onto a C3H background for 4-8 generations before testing). This difference in background could have affected the phenotype of the mice.

Ormestad and colleagues discovered that homozygous Foxf2 knockout mutations cause a severe reduction of type I and type IV collagens and laminin in the intestinal walls of mice. Therefore the investigation of these molecules, and a physical analysis of the guts in the Foxf2^{W174R} homozygotes would be the next logical step in the investigation of Foxf2^{W174R} mortality.

Although the tongue morphology of the Foxf2 null mutant was abnormal [153], no alterations were observed in the level of MyoD – a regulator of muscle differentiation, or myosin heavy chain. The expression of these molecules are therefore unlikely to be altered in the tongues of the Foxf2^{W174R} mutants, which exhibits a generally less severe phenotype than is exhibited by carriers of the null allele and did not suffer from an obvious tongue phenotype. One aspect of the tongue musculature that may however be worth investigating, is the relative density and constituent subunit ratio's of the potassium channels. The tongue is a highly active muscle, and the sporadic nature of its use is likely to incur extremely variable energy demands. There are parallels between the activity of the tongue, and the activity of the heart – which was demonstrated to possess potassium channels that were regulated by FOXF2 [231]. These similarities are not immediately apparent as the heart is required constantly, whereas the tongue is not. However, the energy

demands on the heart can fluctuate massively depending on an organisms activity levels and it during these times of energetic flux that the FOX proteins are thought to play a significant role in the activation of targets that are able to alter the heart's energetic state [231]. Analysis of potassium channels of the intestinal musculature may also be worthwhile for similar reasons.

7.2 Characterisation of the Del36H Eye Phenotype

Heterozygous knockouts of Foxc1 have a different eye phenotype than that of mice with regional monosomy of the Del36H region, so it seems likely that other genes were involved in this phenotype. The possibility that Foxf2 is another contributing factor is of great interest in respect to the relationship between the overlapping function of the genes – and their chromosomal proximity.

The high level of homology that Foxf2 shares with Foxc1 and the similar phenotype seen in the anterior section of the eyes of mouse mutants would suggest that there may be some redundancy of function between the two genes (although not complete redundancy as evidenced by the effect of haploinsufficiency of either gene). It is even possible that Foxc1 and Foxf2 are responsible for the transactivation of an identical array of genes in the eye. It would therefore be interesting to analyse the affect on the eye on a combination of Foxf2 and Foxc1 mutations in compound heterozygotes.

Gould believes that L-Dopa is likely to be important in the formation of the anterior segment of the eye because it is a precursor to catecholamine developmental regulators [194]. As such L-Dopa can be linked to known pathways for many of the genes involved in angle formation such as PAX6, LMX1B, PITX2, PITX3 and for Bmp4 and Tgfb2 which demonstrate similar phenotypes to Foxc1 and Foxc2. PITX2 binds to and inhibits the activity of the FOXC1 protein [127]. Given the similar sequence and structure of these forkhead genes, and the indistinguishable eye phenotype in mutants, it is possible that FOXF2 shares a similar expression profile to FOXC1 in the eye.

One common theme between the observed phenotype of the gut in Foxf2 mutants and the Foxc1 mutant phenotype of the ocular anterior segment is that extracellular matrix is reduced in both systems [154, 278]

Two of the primary structural constituents of ECM are the collagens. Knocking out the Col18a1 gene causes the malformation of the ciliary body and iris [191] and thinning of the cornea and Descemet's membrane is seen in Col8 mutants [192]. These studies indicate that collagen levels are integral to the structure of the anterior segment. Collagen therefore appears to be a downstream target of transcription factors that are associated with anterior segment malformations.

7.3 Alternative Candidates for the Lethal Recessive Screen Mutations

Almost all of the terminal 6p deletions (sub-telomeric deletions) include the forkhead cluster. Given the well characterised and intensely researched eye phenotype associated with Foxc1 it is therefore not surprising that many of the accounts of these patients focus on this gene as the primary causative factor within the region.

However, one patient who carried a micro-deletion in the telomeric region of 6p expressed many of the symptoms of previous sub-telomeric deletions – even though the deletion was only ~200kb in length [279]. This region consisted of only two genes, neither of which were from the forkhead cluster although the deletion breakpoints were within 1.2Mb of FOXC1 (and ~900kb of the most distal member of the FOX cluster - FOXQ1). These were DUSP22 (dual specificity phosphatase) and the most distal 1.4kb of the interferon regulator gene IRF4.

Neither of these genes has previously been associated with the eye and the author does not discuss them other than to mention their absence from the affected chromosome. Instead, the possibility is raised that a long-range positional affect may have disrupted the expression of FOXF2 or FOXC1 [279].

DUSP22 expression in the eye has not been detected, but the eye was not amongst the small number of tissues to undergo expression analysis for this gene [280]. It is a member of the kinase phosphatase gene family. These genes have reported roles in

the regulation of cell proliferation, differentiation and apoptosis via the MAP kinases. Mutations at putative MapK phosphorylation sites of the drosophila gene eyeless (eya) caused the formation of ectopic eyes [281]; eya is now though to be directly downstream of the RTK/RAS/MAPK signalling pathway. In mammals, two of the four homologues of eya, Eya1 and Eya2 are regulated by Pax6 [282].

This information could be used to implicate mouse Dusp22 as another candidate for some of the mutants derived from the lethal recessive screen. No coding mutations were discovered in any of the three forkhead genes despite screening each of the lineages that emerged from the lethal screen for mutations in these genes. There are no immediately obvious alternative candidates in the critical regions for the majority of mutants that demonstrated an eye phenotype. It will therefore be necessary to sequence promoter regions of the forkhead genes, and to also investigate some of the genes without known links to eye phenotypes where there are potential pathways to such an affect.

Geminin is a gene in the Del36H interval with previous linkage to eye development. It is a DNA cell cycle inhibitor that binds and acts antagonistically to the homeobox-containing transcription factor Six3 [283]. Loss of geminin induces the over-proliferation of retinal precursor cells and causes enlarged optic vesicles. However this gene lies between d13Mit15 and d13Mit84, a ~1.3Mb region that has not been identified as a critical region for any of the mutants derived from the lethal recessive screen.

E2f3 is another potential candidate for some of the Del36H mutants with eye phenotypes. It is present in the critical regions of three mutations with structural eye anomalies (412, 624 and 1073); and one other mutation (1616) which did not survive due to the failure of its allantois to fuse. E2f3 is associated with retinoblastoma and its expression can induce lens fibre cells to enter the cell cycle. Cell cycle entry was accompanied by apoptotic cell death resulting in microphthalmia (small eyes) [284].

A prolactin or possibly numerous prolactins may contribute to the Del36H eye phenotype. In humans, prolactin is involved in the activation of ocular blood vessels [189], but a direct comparison of mice and humans is complicated by the large expansion in the mouse prolactin family.

Foxq1 is the third member of the tightly linked chromosome 13 forkhead cluster and its open reading frame was analysed for sequence mutations in each of the mutants by the Ragoussis group. An eye phenotype induced by a Foxq1 mutation would be of particular interest given that Foxc1 and Foxf2 both produce similar ocular phenotypes to each other. There is not yet enough data to consider Foxq1 as a candidate for eye phenotypes but it has been suggested that there may be a relationship between the chromosomal position and temporal order of expression of the genes within the forkhead cluster on mouse chromosome 13 [185]. Microscopic analysis of the iridocorneal angle of satin mice may be worthwhile given that the

phenotype of Foxf2^{W174R} mouse eyes initially went undetected because of their grossly normal appearance.

7.4 Heart Defects

The mutation that was identified in lineage m91 from the lethal recessive screen was localized to a serine to proline alteration in the HMG box domain of Sox4. This mutation caused atrial and ventricular septal defects, in addition to a dysplastic mitral valve [26], which functions to allow blood to pass from the left atrium into the left ventricle. Sox4 homozygous null mutants from an earlier study exhibited aberrant development of the valves and outflow tracts of the heart [285].

There is a high prevalence of heart disorders within the 6p deletion patients – including those with deletions that include Sox4 at 6p22.3 [286], but also in the 6p25 terminal deletions [2, 287].

These aberrant heart features may also be explained by the presence of Foxc1 in region 6p25 as Foxc1 has a role in heart development which it shares with Foxc2. This was highlighted by the analysis of Foxc2 homozygous/Foxc1 heterozygous compound null mutants. Mutant embryos showed an opening between the dorsal aorta and the anterior cardinal vein [120].

7.5 Metabolic Phenotype

Results from the blood and urine metabolic tests gave results that could link the Foxf2 gene to kidney function. Individually these results were tenuous – blood levels of potassium, chloride and calcium in male and urea in female heterozygous mutants were significantly different from wildtypes during initial tests, but secondary testing showed no difference in these biochemical markers. A significant elevation of protein levels was detected in the urine of male heterozygote mutants, although the numbers of mice used in that test were not high enough for a truly representative comparison of this type of ordinal data.

Despite the ambiguity surrounding this data, results of these tests are suggestive of " renal malfunction in the $Foxf2^{W174R}$ mutants. It is a tempting hypothesis that Foxf2 regulates the potassium channels of the kidney by a similar mechanism that which it is proposed to regulate the potassium channels in heart muscle [231]. Potassium is primarily transported passively in the kidney, but active transport is also required for the maintenance of the negative potential of the tubule, the supply of potassium for 2Cl-K and Na-K, and for the secretion of potassium into the cortical collection tube [288].

A kidney phenotype in Foxf2 mutants was unexpected given the apparent absence of FoxF2 from the embryonic and adult kidneys of human and mouse [188, 205] However it is possible that the assays used by Aitola were not sensitive enough to detect the expression of this gene. Furthermore, unpublished work by the

commercial antibody company Abcam uses human kidney cells to demonstrate the effectiveness of their FOXF2 antibody (http://www.abcam.com/index.html?datasheet=23306). The sensitivity of the earlier human expression analysis [205] must also be questioned when FOXF2 has been recently shown to exert an important regulatory effect in cardiac tissue [231], despite the initial northern blot data indicating it was absent from the heart [205].

Alternative hypotheses' to explain the unexpected metabolic screen results could incorporate the potential forkhead binding sites on genes that code for glucose and fatty acid transporters, and enzymes involved in the glycolytic and β-oxidation pathways [231]. Another possibility that could at least explain the elevated potassium levels in $Foxf2^{W174R}$ heterozygotes, is that mutations in SUR1 and KIR6.2 in the pancreas could cause congenital hyperinsulinism [289]. An elevation in insulin levels decreases the excretion of potassium from the blood-stream [290].

Although the lack of a significant difference in the second round of tests casts doubt on the reliability of the initial results, several factors indicated that the explanation may be more involved than a simple type I error. It is unlikely that four of the tested blood biochemicals would all demonstrate a significant difference by chance.

$Foxf2^{W174R}$ mice have shown, in a separate group of experiments, a trend – albeit an insignificant one - towards protection from insulin resistance when fed a high fat diet at 12 weeks of age (M.Goldsworthy, personal communication). However there

was no discernable difference between mutants and wildtypes at 16 weeks. It is therefore feasible that the difference in blood biochemistry results seen at the earlier age may represent a genuine metabolic abnormality that becomes corrected upon homeostatic maturation.

7.6 Startle and Foxf2

There was a clear difference between phenotypes in wild-type and mutants in the latest and largest startle dataset. The reliability of the earlier tests of both Startle and PPI could have been adversely affected by a range of factors that may have confounded the data. For example the sound quality in the startle boxes was poor, there was a higher degree of background strain variability, and the environment was less well regulated before the animals were moved into individually ventilated cages. Strain background is particularly important to consider due to the elevated startle response of the BALB/c strain [254]which was used for initial mutagenesis of the mice that were included in the Harwell ENU archives. Startle results from these data therefore indicate that the Foxf2^{W174R} mutation affects the anxiety pathway

The 6p deletion syndromes are associated with a range of neurological defects – in common with many other large chromosomal deletions, 6p deletions often cause mental retardation [2, 73, 75, 287]. There are also a range of loci for specific psychological disorders such as schizophrenia [216], attention hyperactivity disorder (which co-segregates with dyslexia) [291], anxiety and depression [292] that have been localised to 6p. These types of neurological disorder would be difficult or

impossible to assess directly in mouse models. However, they are associated with intermediate traits that can allow the measurement of a quantifiable phenotype [293].

The acoustic startle response (ASR) is one such intermediate trait. It can provide a measure of anxiety – a neurobehavioural deficit which can show co-morbidity with depression [292]. The elevated startle response demonstrated in the Foxf2^{W174R} mutant could therefore suggest that Foxf2 could in future be regarded as a candidate for these disorders. However none of the 6p loci for the human disorders that have been identified so far have been localised the 6p25 – there is a schizophrenia locus at 6p23 and 6p21, attention hyperactivity disorder at 6p [291] and a potential anxiety/depression locus at 6p21.2.

7.7 Molecular Phenotyping

Evidence obtained during the phenotyping aspect of the project would suggest that Foxf2^{W174R} is a hypomorphic mutation and the molecular evidence supports this. The EMSA profile of the FOXF2^{W174R} mutagenised plasmid is of a greatly reduced intensity compared to the other samples, but there is a faint band at the expected position which is absent from the negative controls. This suggests that Foxf2^{W174R} is a hypomorphic rather than a null mutation.

Although the reporter that contained FOXF2 binding sites did not appear to be activated by the peptide product, there is evidence to suggest that FOXF2 activates the empty reporter. As such, despite the lack of a robust experimental model, some

information can be obtained from these data if the assumption that the endogenous activation is a real effect.

There was no significant difference between the cells transfected with the FOXF2^{W174R} plasmid and the empty expression vector. Because of the phenotypic and EMSA evidence suggesting that this mutation is a hypomorphic allele however, it seems likely that the levels of activation may have been too low to be detected in the FOXF2^{W174R} plasmid samples.

The V410F mutant plasmid showed no difference to the wildtype plasmid in the EMSA, so DNA binding appears to occur as normal in this mutant. In the activation assay, the V410F samples were elevated significantly above the empty expression vector samples, and were significantly reduced in comparison with the wildtype samples – a reduction that was significant in two of three timepoints. This implies that the V410F mutation causes a slight reduction in the trans-activational properties of the FOXF2 peptide – if the assumption is made that there is genuine activation of the endogenous sites on pTAL-SEAP by the FOXF2 fusion peptide.

This data must be viewed with caution due to the apparent failure of the reporter containing the inserted binding site. These results are not robust enough to support any concrete theories regarding the activities of the different FOXF2 isoforms and these experiments will need to be repeated with a redesigned reporter, perhaps with the vector pmluc3-cmv, as used in the last EMSA experiment.

7.8 Future Work

If a repeat of the transactivation assay supported the observation that the FOXF2V410F peptide operated at a slightly reduced efficiency, an investigation of the Phenotype of the Foxf2V412F mouse could build on the information ascertained from this study. The FOXF2V412F mutation is in the same activation domain as the polyglycine expansion that is suspected to cause the glaucoma in the human family identified by H. Vieira – so this would be of particular interest to analyse further. One potentially exciting prospect with regard to the V412F mutation is that it is possible that these mutations will cause a more subtle effect on the function of the protein, that may allow the survival of homozygous mutants.

A homozygous mutation with a phenotype that is less severe than $Foxf2^{W174R}$ could be very informative for the analysis of the function of the gene in the tissues of older mice. An adult phenotype for a homozygous Foxf2 mutation has not yet been possible because the mutations that have been analysed to date all cause neonatal lethality, although the severity of the $Foxf2^{W174R}$ mutation appears to be reduced in comparison with the null allele. For this reason, it is definitely worth investigating the potential of both the V412F and the splice site mutation.

The key first step in the analysis of the splice site mutant, would be the determination of expression of the full mRNA in these mutants. A comparison between RNA taken from Foxf2 expressing tissues of splice site mutants and

wildtype siblings could be achieved by real-time PCR. Primers should span both exons.

Previous domain studies have identified the region responsible for TBP binding in human FOXF2 [242]. The region as between 146-166 of the FOXF2 construct, which is at amino acid positions 182-202 of the full peptide sequence as it is currently published. This is only separated from the W174R mutation by 7 amino acids. It is therefore possible that the W174R mutation may interfere with the interaction between FOXF2 and TBP, in addition to reducing DNA binding affinity.

The weight of phenotypic and molecular evidence would suggest that W174R is a hypomorphic mutation. Therefore it may be worth investigating to see if an inhibited TBP interaction alleviates or enhances the affect on the peptide function that is caused by its reduced DNA affinity.

Another potential effect of the mutation is a reduction in the nuclear localisation signal on the Foxf2 protein because this has been localised to the forkhead domain [242]. It is possible that the mutant band in the EMSA was weak because the protein was not concentrated in the nuclear fraction. An additional experiment to answer this question would involve repeating the experiment with the cytoplasmic fraction of the transfected cells. If the $Foxf2^{W174R}$ mutation had affected nuclear localisation but not DNA binding, a strong band would be observed in the $Foxf2^{W174R}$ signal which would be absent from the other samples.

References

1. Ji, Y., et al., *The ancestral gene for transcribed, low-copy repeats in the Prader-Willi/Angelman region encodes a large protein implicated in protein trafficking, which is deficient in mice with neuromuscular and spermiogenic abnormalities.* Hum Mol Genet, 1999. **8**(3): p. 533-42.
2. Mirza, G., et al., *Refined genotype-phenotype correlations in cases of chromosome 6p deletion syndromes.* Eur J Hum Genet, 2004. **12**(9): p. 718-28.
3. Cerruti Mainardi, P., *Cri du Chat syndrome.* Orphanet J Rare Dis, 2006. **1**: p. 33.
4. Abeysinghe, S.S., et al., *Translocation and gross deletion breakpoints in human inherited disease and cancer I: Nucleotide composition and recombination-associated motifs.* Hum Mutat, 2003. **22**(3): p. 229-44.
5. Miller, O.J. and W.R. Breg, *Autosomal chromosome disorders and variations.* N Engl J Med, 1976. **294**(11): p. 596-8.
6. Marchetti, F. and A.J. Wyrobek, *Mechanisms and consequences of paternally-transmitted chromosomal abnormalities.* Birth Defects Res C Embryo Today, 2005. **75**(2): p. 112-29.
7. Bruder, C.E., et al., *High resolution deletion analysis of constitutional DNA from neurofibromatosis type 2 (NF2) patients using microarray-CGH.* Hum Mol Genet, 2001. **10**(3): p. 271-82.
8. Schmickel, R.D., *Contiguous gene syndromes: a component of recognizable syndromes.* J Pediatr, 1986. **109**(2): p. 231-41.
9. Budarf, M.L. and B.S. Emanuel, *Progress in the autosomal segmental aneusomy syndromes (SASs): single or multi-locus disorders?* Hum Mol Genet, 1997. **6**(10): p. 1657-65.
10. Lehmann, O.J., et al., *Ocular developmental abnormalities and glaucoma associated with interstitial 6p25 duplications and deletions.* Invest Ophthalmol Vis Sci, 2002. **43**(6): p. 1843-9.
11. Barber, J.C., et al., *Interstitial deletions without phenotypic effect: prenatal diagnosis of a new family and brief review.* Prenat Diagn, 1991. **11**(6): p. 411-6.
12. Eichler, E.E., et al., *Duplication of a gene-rich cluster between 16p11.1 and Xq28: a novel pericentromeric-directed mechanism for paralogous genome evolution.* Hum Mol Genet, 1996. **5**(7): p. 899-912.
13. Eichler, E.E., *Masquerading repeats: paralogous pitfalls of the human genome.* Genome Res, 1998. **8**(8): p. 758-62.
14. Callinan, P.A., et al., *Alu retrotransposition-mediated deletion.* J Mol Biol, 2005. **348**(4): p. 791-800.
15. Bailey, J.A. and E.E. Eichler, *Primate segmental duplications: crucibles of evolution, diversity and disease.* Nat Rev Genet, 2006. **7**(7): p. 552-64.
16. Shaw, C.J. and J.R. Lupski, *Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease.* Hum Mol Genet, 2004. **13 Spec No 1**: p. R57-64.

17. Shaw, C.J. and J.R. Lupski, *Non-recurrent 17p11.2 deletions are generated by homologous and non-homologous mechanisms.* Hum Genet, 2005. **116**(1-2): p. 1-7.
18. Lupski, J.R., *Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits.* Trends Genet, 1998. **14**(10): p. 417-22.
19. Stankiewicz, P. and J.R. Lupski, *Genome architecture, rearrangements and genomic disorders.* Trends Genet, 2002. **18**(2): p. 74-82.
20. Cattanach, B.M., et al., *Large deletions and other gross forms of chromosome imbalance compatible with viability and fertility in the mouse.* Nat Genet, 1993a. **3**(1): p. 56-61.
21. Arkell, R.M., et al., *Genetic, physical, and phenotypic characterization of the Del(13)Svea36H mouse.* Mamm Genome, 2001. **12**(9): p. 687-94.
22. Rinchik, E.M., D.A. Carpenter, and P.B. Selby, *A strategy for fine-structure functional analysis of a 6- to 11-centimorgan region of mouse chromosome 7 by high-efficiency mutagenesis.* Proc Natl Acad Sci U S A, 1990. **87**(3): p. 896-900.
23. Gausz, J., et al., *Genetic characterization of the region between 86F1,2 and 87B15 on chromosome 3 of Drosophila melanogaster.* Genetics, 1981. **98**(4): p. 775-89.
24. Schimenti, J. and M. Bucan, *Functional genomics in the mouse: phenotype-based mutagenesis screens.* Genome Res, 1998. **8**(7): p. 698-710.
25. Walz, K., P. Fonseca, and J.R. Lupski, *Animal Models for Human Contiguous Gene Syndromes and Other Genomic Disorders.* Genetics and Molecular Biology, 2004. **27**(3): p. 305-320.
26. Bogani, D., et al., *Dissecting the genetic complexity of human 6p deletion syndromes by using a region-specific, phenotype-driven mouse screen.* Proc Natl Acad Sci U S A, 2005. **102**(35): p. 12477-82.
27. Ramirez-Solis, R., P. Liu, and A. Bradley, *Chromosome engineering in mice.* Nature, 1995. **378**(6558): p. 720-4.
28. Forlino, A., et al., *Use of the Cre/lox recombination system to develop a non-lethal knock-in murine model for osteogenesis imperfecta with an alpha1(I) G349C substitution. Variability in phenotype in BrtlIV mice.* J Biol Chem, 1999. **274**(53): p. 37923-31.
29. Herault, Y., et al., *Engineering chromosomes in mice through targeted meiotic recombination (TAMERE).* Nat Genet, 1998. **20**(4): p. 381-4.
30. Russell, W.L., *X-ray-induced mutations in mice.* Cold Spring Harb Symp Quant Biol, 1951. **16**: p. 327-36.
31. Davis, A.P. and M.J. Justice, *An Oak Ridge legacy: the specific locus test and its role in mouse mutagenesis.* Genetics, 1998. **148**(1): p. 7-12.
32. Russell, L.B., *Definition of functional units in a small chromosomal segment of the mouse and its use in interpreting the nature of radiation-induced mutations.* Mutat Res, 1971. **11**(1): p. 107-23.
33. Rinchik, E.M., et al., *Molecular genetic analysis of the dilute-short ear (d-se) region of the mouse.* Genetics, 1986. **112**(2): p. 321-42.

34. Rinchik, E.M., *Molecular analysis of heritable mouse mutations.* Environ Health Perspect, 1987. **74**: p. 41-8.
35. Dietrich, W.F., et al., *A comprehensive genetic map of the mouse genome.* Nature, 1996. **380**(6570): p. 149-52.
36. Hudson, T.J., et al., *A radiation hybrid map of mouse genes.* Nat Genet, 2001. **29**(2): p. 201-5.
37. Nusbaum, C., et al., *A YAC-based physical map of the mouse genome.* Nat Genet, 1999. **22**(4): p. 388-93.
38. Burke, D.T., G.F. Carle, and M.V. Olson, *Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors.* Science, 1987. **236**(4803): p. 806-12.
39. Kelavkar, U. and K. Shah, *Advances in the Human Genome Project. A review.* Mol Biol Rep, 1998. **25**(1): p. 27-43.
40. McPherson, J.D., *Sequence ready-or not?* Genome Res, 1997. **7**(12): p. 1111-3.
41. Ioannou, P.A., et al., *A new bacteriophage P1-derived vector for the propagation of large human DNA fragments.* Nat Genet, 1994. **6**(1): p. 84-9.
42. Shizuya, H., et al., *Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector.* Proc Natl Acad Sci U S A, 1992. **89**(18): p. 8794-7.
43. Waterston, R.H., et al., *Initial sequencing and comparative analysis of the mouse genome.* Nature, 2002. **420**(6915): p. 520-62.
44. Venter, J.C., et al., *Shotgun sequencing of the human genome.* Science, 1998. **280**(5369): p. 1540-2.
45. Green, E.D., *Strategies for the systematic sequencing of complex genomes.* Nat Rev Genet, 2001. **2**(8): p. 573-83.
46. Russell, W.L., et al., *Specific-locus test shows ethylnitrosourea to be the most potent mutagen in the mouse.* Proc Natl Acad Sci U S A, 1979. **76**(11): p. 5818-9.
47. Popp, R.A., et al., *Analysis of a mouse alpha-globin gene mutation induced by ethylnitrosourea.* Genetics, 1983. **105**(1): p. 157-67.
48. Cordes, S.P., *N-ethyl-N-nitrosourea mutagenesis: boarding the mouse mutant express.* Microbiol Mol Biol Rev, 2005. **69**(3): p. 426-39.
49. Justice, M.J., et al., *Mouse ENU mutagenesis.* Hum Mol Genet, 1999. **8**(10): p. 1955-63.
50. Brown, S.D. and J. Peters, *Combining mutagenesis and genomics in the mouse--closing the phenotype gap.* Trends Genet, 1996. **12**(11): p. 433-5.
51. Noveroske, J.K., J.S. Weber, and M.J. Justice, *The mutagenic action of N-ethyl-N-nitrosourea in the mouse.* Mamm Genome, 2000. **11**(7): p. 478-83.
52. Hansen, J., et al., *A large-scale, gene-driven mutagenesis approach for the functional analysis of the mouse genome.* Proc Natl Acad Sci U S A, 2003. **100**(17): p. 9918-22.
53. Wilkie, A.O., *The molecular basis of genetic dominance.* J Med Genet, 1994. **31**(2): p. 89-98.
54. Vitaterna, M.H., et al., *Mutagenesis and mapping of a mouse gene, Clock, essential for circadian behavior.* Science, 1994. **264**(5159): p. 719-25.

55. Nolan, P.M., et al., *A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse.* Nat Genet, 2000. **25**(4): p. 440-3.
56. Hrabe de Angelis, M.H., et al., *Genome-wide, large-scale production of mutant mice by ENU mutagenesis.* Nat Genet, 2000. **25**(4): p. 444-7.
57. Rogers, D.C., et al., *Behavioral and functional analysis of mouse phenotype: SHIRPA, a proposed protocol for comprehensive phenotype assessment.* Mamm Genome, 1997. **8**(10): p. 711-3.
58. Nolan, P.M., et al., *Implementation of a large-scale ENU mutagenesis program: towards increasing the mouse mutant resource.* Mamm Genome, 2000. **11**(7): p. 500-6.
59. Gerits, N., W. Van Belle, and U. Moens, *Transgenic mice expressing constitutive active MAPKAPK5 display gender-dependent differences in exploration and activity.* Behav Brain Funct, 2007. **3**: p. 58.
60. Bogue, M., *Mouse Phenome Project: understanding human biology through mouse genetics and genomics.* J Appl Physiol, 2003. **95**(4): p. 1335-7.
61. Bogue, M.A. and S.C. Grubb, *The Mouse Phenome Project.* Genetica, 2004. **122**(1): p. 71-4.
62. Holmes, A., et al., *Behavioral profiles of inbred strains on novel olfactory, spatial and emotional tests for reference memory in mice.* Genes Brain Behav, 2002. **1**(1): p. 55-69.
63. Bolivar, V.J., et al., *Habituation of activity in an open field: A survey of inbred strains and F1 hybrids.* Behav Genet, 2000. **30**(4): p. 285-93.
64. Crabbe, J.C., D. Wahlsten, and B.C. Dudek, *Genetics of mouse behavior: interactions with laboratory environment.* Science, 1999. **284**(5420): p. 1670-2.
65. Crusio, W.E., *Flanking gene and genetic background problems in genetically manipulated mice.* Biol Psychiatry, 2004. **56**(6): p. 381-5.
66. Champy, M.F., et al., *Mouse functional genomics requires standardization of mouse handling and housing conditions.* Mamm Genome, 2004. **15**(10): p. 768-83.
67. Tucci, V., et al., *Gene-environment interactions differentially affect mouse strain behavioral parameters.* Mamm Genome, 2006. **17**(11): p. 1113-20.
68. Chesler, E.J., et al., *Influences of laboratory environment on behavior.* Nat Neurosci, 2002. **5**(11): p. 1101-2.
69. Brown, S.D., P. Chambon, and M.H. de Angelis, *EMPReSS: standardized phenotype screens for functional annotation of the mouse genome.* Nat Genet, 2005. **37**(11): p. 1155.
70. Brown, S.D., J.M. Hancock, and H. Gates, *Understanding mammalian genetic systems: the challenge of phenotyping in the mouse.* PLoS Genet, 2006. **2**(8): p. e118.
71. Caluseriu, O., et al., *Schizophrenia in an adult with 6p25 deletion syndrome.* Am J Med Genet A, 2006. **140**(11): p. 1208-13.
72. Chen, K.M., et al., *Mild developmental delay in terminal chromosome 6p deletion.* Am J Med Genet A, 2004. **129A**(2): p. 201-5.

73. Descipio, C., et al., *Subtelomeric deletions of chromosome 6p: molecular and cytogenetic characterization of three new cases with phenotypic overlap with Ritscher-Schinzel (3C) syndrome.* Am J Med Genet A, 2005. **134A**(1): p. 3-11.
74. Klein, O.D., et al., *Case report: Y;6 translocation with deletion of 6p.* Clin Dysmorphol, 2005. **14**(2): p. 93-6.
75. Le Caignec, C., et al., *Subtelomeric 6p deletion: clinical, FISH, and array CGH characterization of two cases.* Am J Med Genet A, 2005. **132A**(2): p. 175-80.
76. Maclean, K., et al., *Axenfeld-Rieger malformation and distinctive facial features: Clues to a recognizable 6p25 microdeletion syndrome.* Am J Med Genet A, 2005. **132**(4): p. 381-5.
77. Martinez-Glez, V., et al., *Clinical presentation of a variant of Axenfeld-Rieger syndrome associated with subtelomeric 6p deletion.* Eur J Med Genet, 2007. **50**(2): p. 120-7.
78. Lin, R.J., et al., *Terminal deletion of 6p results in a recognizable phenotype.* Am J Med Genet A, 2005. **136**(2): p. 162-8.
79. Davies, A.F., et al., *Delineation of two distinct 6p deletion syndromes.* Hum Genet, 1999. **104**(1): p. 64-72.
80. Gould, D.B., et al., *Phenotypic and molecular assessment of seven patients with 6p25 deletion syndrome: relevance to ocular dysgenesis and hearing impairment.* BMC Med Genet, 2004. **5**: p. 17.
81. Kume, T., et al., *The forkhead/winged helix gene Mfl is disrupted in the pleiotropic mouse mutation congenital hydrocephalus.* Cell, 1998. **93**(6): p. 985-96.
82. Tease, C. and G. Fisher, *Cytogenetic detection of four new, viable deletions in the progeny of X-irradiated females.* Mouse Genome, 1993. **91**: p. 855.
83. Rhodes, M., et al., *A high-resolution microsatellite map of the mouse genome.* Genome Res, 1998. **8**(5): p. 531-42.
84. Mallon, A.M., et al., *Organization and evolution of a gene-rich region of the mouse genome: a 12.7-Mb region deleted in the Del(13)Svea36H mouse.* Genome Res, 2004. **14**(10A): p. 1888-901.
85. Bargmann, C.I., *Olfactory receptors, vomeronasal receptors, and the organization of olfactory information.* Cell, 1997. **90**(4): p. 585-7.
86. Keverne, E.B., *The vomeronasal organ.* Science, 1999. **286**(5440): p. 716-20.
87. Wiemers, D.O., et al., *The mouse prolactin gene family locus.* Endocrinology, 2003. **144**(1): p. 313-25.
88. Ogg, S.L., et al., *Expression of butyrophilin (Btn1a1) in lactating mammary gland is essential for the regulated secretion of milk-lipid droplets.* Proc Natl Acad Sci U S A, 2004. **101**(27): p. 10084-9.
89. Silverman, G.A., et al., *The serpins are an expanding superfamily of structurally similar but functionally diverse proteins. Evolution, mechanism of inhibition, novel functions, and a revised nomenclature.* J Biol Chem, 2001. **276**(36): p. 33293-6.

90. Kaiserman, D., et al., *Comparison of human chromosome 6p25 with mouse chromosome 13 reveals a greatly expanded ov-serpin gene repertoire in the mouse.* Genomics, 2002. **79**(3): p. 349-62.
91. Jürgens, G. and D. Weigel, *Terminal versus segmental development in the Drosophila embryo: the role of the homeotic gene fork head.* Development Genes and Evolution, 1988. **197**(6): p. 345-354.
92. Lai, E., et al., *HNF-3A, a hepatocyte-enriched transcription factor of novel structure is regulated transcriptionally.* Genes Dev, 1990. **4**(8): p. 1427-36.
93. Mazet, F., et al., *Phylogenetic relationships of the Fox (Forkhead) gene family in the Bilateria.* Gene, 2003. **316**: p. 79-89.
94. Carlsson, P. and M. Mahlapuu, *Forkhead transcription factors: key players in development and metabolism.* Dev Biol, 2002. **250**(1): p. 1-23.
95. Clark, K.L., et al., *Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5.* Nature, 1993. **364**(6436): p. 412-20.
96. Marsden, I., C. Jin, and X. Liao, *Structural changes in the region directly adjacent to the DNA-binding helix highlight a possible mechanism to explain the observed changes in the sequence-specific binding of winged helix proteins.* J Mol Biol, 1998. **278**(2): p. 293-9.
97. van Dongen, M.J., et al., *Solution structure and dynamics of the DNA-binding domain of the adipocyte-transcription factor FREAC-11.* J Mol Biol, 2000. **296**(2): p. 351-9.
98. Weigelt, J., et al., *Solution structure of the DNA binding domain of the human forkhead transcription factor AFX (FOXO4).* Biochemistry, 2001. **40**(20): p. 5861-9.
99. Cirillo, L.A. and K.S. Zaret, *Specific interactions of the wing domains of FOXA1 transcription factor with DNA.* J Mol Biol, 2007. **366**(3): p. 720-4.
100. Cirillo, L.A. and K.S. Zaret, *An early developmental transcription factor complex that is more stable on nucleosome core particles than on free DNA.* Mol Cell, 1999. **4**(6): p. 961-9.
101. Cirillo, L.A., et al., *Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4.* Mol Cell, 2002. **9**(2): p. 279-89.
102. Wang, I.C., et al., *Forkhead box M1 regulates the transcriptional network of genes essential for mitotic progression and genes encoding the SCF (Skp2-Cks1) ubiquitin ligase.* Mol Cell Biol, 2005. **25**(24): p. 10875-94.
103. Hawke, T.J., N. Jiang, and D.J. Garry, *Absence of p21CIP rescues myogenic progenitor cell proliferative and regenerative capacity in Foxk1 null mice.* J Biol Chem, 2003. **278**(6): p. 4015-20.
104. Garry, D.J., et al., *Persistent expression of MNF identifies myogenic stem cells in postnatal muscles.* Dev Biol, 1997. **188**(2): p. 280-94.
105. Seoane, J., et al., *Integration of Smad and forkhead pathways in the control of neuroepithelial and glioblastoma cell proliferation.* Cell, 2004. **117**(2): p. 211-23.
106. Dijkers, P.F., et al., *Forkhead transcription factor FKHR-L1 modulates cytokine-dependent transcriptional regulation of p27(KIP1).* Mol Cell Biol, 2000. **20**(24): p. 9138-48.

107. Medema, R.H., et al., *AFX-like Forkhead transcription factors mediate cell-cycle regulation by Ras and PKB through p27kip1.* Nature, 2000. **404**(6779): p. 782-7.
108. Ang, S.L. and J. Rossant, *HNF-3 beta is essential for node and notochord formation in mouse development.* Cell, 1994. **78**(4): p. 561-74.
109. Friedman, J.R. and K.H. Kaestner, *The Foxa family of transcription factors in development and metabolism.* Cell Mol Life Sci, 2006. **63**(19-20): p. 2317-28.
110. Hong, H.K., J.H. Lass, and A. Chakravarti, *Pleiotropic skeletal and ocular phenotypes of the mouse mutation congenital hydrocephalus (ch/Mf1) arise from a winged helix/forkhead transcriptionfactor gene.* Hum Mol Genet, 1999. **8**(4): p. 625-37.
111. Sasaki, H. and B.L. Hogan, *Differential expression of multiple fork head related genes during gastrulation and axial pattern formation in the mouse embryo.* Development, 1993. **118**(1): p. 47-59.
112. Jeong, J., et al., *Hedgehog signaling in the neural crest cells regulates the patterning and growth of facial primordia.* Genes Dev, 2004. **18**(8): p. 937-51.
113. Wu, S.C., et al., *Mouse Mesenchyme forkhead 2 (Mf2): expression, DNA binding and induction by sonic hedgehog during somitogenesis.* Mech Dev, 1998. **70**(1-2): p. 3-13.
114. Furumoto, T.A., et al., *Notochord-dependent expression of MFH1 and PAX1 cooperates to maintain the proliferation of sclerotome cells during the vertebral column development.* Dev Biol, 1999. **210**(1): p. 15-29.
115. Ingram, W.J., et al., *Novel genes regulated by Sonic Hedgehog in pluripotent mesenchymal cells.* Oncogene, 2002. **21**(53): p. 8196-205.
116. Mahlapuu, M., S. Enerback, and P. Carlsson, *Haploinsufficiency of the forkhead gene Foxf1, a target for sonic hedgehog signaling, causes lung and foregut malformations.* Development, 2001. **128**(12): p. 2397-406.
117. Yamagishi, H., et al., *Tbx1 is regulated by tissue-specific forkhead proteins through a common Sonic hedgehog-responsive enhancer.* Genes Dev, 2003. **17**(2): p. 269-81.
118. Wan, H., et al., *Compensatory roles of Foxa1 and Foxa2 during lung morphogenesis.* J Biol Chem, 2005. **280**(14): p. 13809-16.
119. Wang, B., et al., *Foxp1 regulates cardiac outflow tract, endocardial cushion morphogenesis and myocyte proliferation and maturation.* Development, 2004. **131**(18): p. 4477-87.
120. Seo, S., et al., *The forkhead transcription factors, Foxc1 and Foxc2, are required for arterial specification and lymphatic sprouting during vascular development.* Dev Biol, 2006. **294**(2): p. 458-70.
121. Shu, W., et al., *Characterization of a new subfamily of winged-helix/forkhead (Fox) genes that are expressed in the lung and act as transcriptional repressors.* J Biol Chem, 2001. **276**(29): p. 27488-97.
122. Chi, N.C., et al., *Foxn4 directly regulates tbx2b expression and atrioventricular canal formation.* Genes Dev, 2008. **22**(6): p. 734-9.

123. Czech, M.P., *Insulin's expanding control of forkheads.* Proc Natl Acad Sci U S A, 2003. **100**(20): p. 11198-200.
124. Lai, C.S., et al., *A forkhead-domain gene is mutated in a severe speech and language disorder.* Nature, 2001. **413**(6855): p. 519-23.
125. Fontenot, J.D., M.A. Gavin, and A.Y. Rudensky, *Foxp3 programs the development and function of CD4+CD25+ regulatory T cells.* Nat Immunol, 2003. **4**(4): p. 330-6.
126. Lehmann, O.J., et al., *Fox's in development and disease.* Trends Genet, 2003. **19**(6): p. 339-44.
127. Berry, F.B., et al., *Functional interactions between FOXC1 and PITX2 underlie the sensitivity to FOXC1 gene dose in Axenfeld-Rieger syndrome and anterior segment dysgenesis.* Hum Mol Genet, 2006. **15**(6): p. 905-19.
128. Lehmann, O.J., et al., *Novel anterior segment phenotypes resulting from forkhead gene alterations: evidence for cross-species conservation of function.* Invest Ophthalmol Vis Sci, 2003. **44**(6): p. 2627-33.
129. Osborne, L.R., *'Forkhead' gene expression balanced on a knife-edge.* Trends Mol Med, 2001. **7**(2): p. 51.
130. Nishimura, D.Y., et al., *The forkhead transcription factor gene FKHL7 is responsible for glaucoma phenotypes which map to 6p25.* Nat Genet, 1998. **19**(2): p. 140-7.
131. Mears, A.J., et al., *Mutations of the forkhead/winged-helix gene, FKHL7, in patients with Axenfeld-Rieger anomaly.* Am J Hum Genet, 1998. **63**(5): p. 1316-28.
132. Fang, J., et al., *Mutations in FOXC2 (MFH-1), a forkhead family transcription factor, are responsible for the hereditary lymphedema-distichiasis syndrome.* Am J Hum Genet, 2000. **67**(6): p. 1382-8.
133. De Baere, E., et al., *Spectrum of FOXL2 gene mutations in blepharophimosis-ptosis-epicanthus inversus (BPES) families demonstrates a genotype--phenotype correlation.* Hum Mol Genet, 2001. **10**(15): p. 1591-600.
134. Kaestner, K.H., et al., *Inactivation of the winged helix transcription factor HNF3alpha affects glucose homeostasis and islet glucagon gene expression in vivo.* Genes Dev, 1999. **13**(4): p. 495-504.
135. Sund, N.J., et al., *Tissue-specific deletion of Foxa2 in pancreatic beta cells results in hyperinsulinemic hypoglycemia.* Genes Dev, 2001. **15**(13): p. 1706-15.
136. Shen, W., et al., *Foxa3 (hepatocyte nuclear factor 3gamma) is required for the regulation of hepatic GLUT2 expression and the maintenance of glucose homeostasis during a prolonged fast.* J Biol Chem, 2001. **276**(46): p. 42812-7.
137. Wehr, R., et al., *Fkh5-deficient mice show dysgenesis in the caudal midbrain and hypothalamic mammillary body.* Development, 1997. **124**(22): p. 4447-56.
138. Smith, R.S., et al., *Haploinsufficiency of the transcription factors FOXC1 and FOXC2 results in aberrant ocular development.* Hum Mol Genet, 2000. **9**(7): p. 1021-32.

139. Lehmann, O.J., et al., *Chromosomal duplication involving the forkhead transcription factor gene FOXC1 causes iris hypoplasia and glaucoma.* Am J Hum Genet, 2000. **67**(5): p. 1129-35.
140. Iida, K., et al., *Essential roles of the winged helix transcription factor MFH-1 in aortic arch patterning and skeletogenesis.* Development, 1997. **124**(22): p. 4627-38.
141. Hatini, V., et al., *Essential role of stromal mesenchyme in kidney morphogenesis revealed by targeted disruption of Winged Helix transcription factor BF-2.* Genes Dev, 1996. **10**(12): p. 1467-78.
142. Kume, T., K. Deng, and B.L. Hogan, *Minimal phenotype of mice homozygous for a null mutation in the forkhead/winged helix gene, Mf2.* Mol Cell Biol, 2000. **20**(4): p. 1419-25.
143. Hanna, L.A., et al., *Requirement for Foxd3 in maintaining pluripotent cells of the early mouse embryo.* Genes Dev, 2002. **16**(20): p. 2650-61.
144. Tompers, D.M., et al., *Foxd3 is required in the trophoblast progenitor cell lineage of the mouse embryo.* Dev Biol, 2005. **285**(1): p. 126-37.
145. Minoretti, P., et al., *A W148R mutation in the human FOXD4 gene segregating with dilated cardiomyopathy, obsessive-compulsive disorder, and suicidality.* Int J Mol Med, 2007. **19**(3): p. 369-72.
146. Yan, B., K.M. Neilson, and S.A. Moody, *foxD5 plays a critical upstream role in regulating neural ectodermal fate and the onset of neural differentiation.* Dev Biol, 2009.
147. De Felice, M., et al., *A mouse model for hereditary thyroid dysgenesis and cleft palate.* Nat Genet, 1998. **19**(4): p. 395-8.
148. Clifton-Bligh, R.J., et al., *Mutation of the gene encoding human TTF-2 associated with thyroid agenesis, cleft palate and choanal atresia.* Nat Genet, 1998. **19**(4): p. 399-401.
149. Blixt, A., et al., *A forkhead gene, FoxE3, is essential for lens epithelial proliferation and closure of the lens vesicle.* Genes Dev, 2000. **14**(2): p. 245-54.
150. Brownell, I., M. Dirksen, and M. Jamrich, *Forkhead Foxe3 maps to the dysgenetic lens locus and is critical in lens development and differentiation.* Genesis, 2000. **27**(2): p. 81-93.
151. Semina, E.V., et al., *Mutations in the human forkhead transcription factor FOXE3 associated with anterior segment ocular dysgenesis and cataracts.* Hum Mol Genet, 2001. **10**(3): p. 231-6.
152. Kalinichenko, V.V., et al., *Defects in pulmonary vasculature and perinatal lung hemorrhage in mice heterozygous null for the Forkhead Box f1 transcription factor.* Dev Biol, 2001. **235**(2): p. 489-506.
153. Wang, T., et al., *Forkhead transcription factor Foxf2 (LUN)-deficient mice exhibit abnormal development of secondary palate.* Dev Biol, 2003. **259**(1): p. 83-94.
154. Ormestad, M., et al., *Foxf1 and Foxf2 control murine gut development by limiting mesenchymal Wnt signaling and promoting extracellular matrix production.* Development, 2006. **133**(5): p. 833-43.

155. Vieira, H., et al., *Mutation of the forkhead gene FOXF2 causes iris hypoplasia and glaucoma*, in *Abstracts of papers presented at the thirteenth Genetics Society's Mammalian Genetics and Development Workshop held at the Institute of Child Health, University College London on 27 and 28 November 2002*, A. Copp and E. Fisher, Editors. 2002, Cambridge University Press: Genet. Res., Camb. p. 231.
156. Xuan, S., et al., *Winged helix transcription factor BF-1 is essential for the development of the cerebral hemispheres.* Neuron, 1995. **14**(6): p. 1141-52.
157. Hoodless, P.A., et al., *FoxH1 (Fast) functions to specify the anterior primitive streak in the mouse.* Genes Dev, 2001. **15**(10): p. 1257-71.
158. Yamamoto, M., et al., *The transcription factor FoxH1 (FAST) mediates Nodal signaling during anterior-posterior patterning and node formation in the mouse.* Genes Dev, 2001. **15**(10): p. 1242-56.
159. Hulander, M., et al., *The winged helix transcription factor Fkh10 is required for normal development of the inner ear.* Nat Genet, 1998. **20**(4): p. 374-6.
160. Brody, S.L., et al., *Ciliogenesis and left-right axis defects in forkhead factor HFH-4-null mice.* Am J Respir Cell Mol Biol, 2000. **23**(1): p. 45-51.
161. Hannenhalli, S., et al., *Transcriptional genomics associates FOX transcription factors with human heart failure.* Circulation, 2006. **114**(12): p. 1269-76.
162. Alexander, M.S., *The roles of forkhead transcription factors in stem cells and myogenesis*. 2007, The University of Texas.
163. Garry, D.J., et al., *Myogenic stem cell function is impaired in mice lacking the forkhead/winged helix protein MNF.* Proc Natl Acad Sci U S A, 2000. **97**(10): p. 5416-21.
164. Wijchers, P.J.E.C., *Forkhead transcription factors in brain development : 'Fox hunting in midbrain dopaminergic neurons'*. 2005, Universiteit Utrecht.
165. Kaestner, K.H., et al., *The mesenchymal winged helix transcription factor Fkh6 is required for the control of gastrointestinal proliferation and differentiation.* Genes Dev, 1997. **11**(12): p. 1583-95.
166. Perreault, N., et al., *Foxl1 controls the Wnt/beta-catenin pathway by modulating the expression of proteoglycans in the gut.* J Biol Chem, 2001. **276**(46): p. 43328-33.
167. Crisponi, L., et al., *The putative forkhead transcription factor FOXL2 is mutated in blepharophimosis/ptosis/epicanthus inversus syndrome.* Nat Genet, 2001. **27**(2): p. 159-66.
168. Korver, W., et al., *Uncoupling of S phase and mitosis in cardiomyocytes and hepatocytes lacking the winged-helix transcription factor Trident.* Curr Biol, 1998. **8**(24): p. 1327-30.
169. Kim, I.M., et al., *The forkhead box m1 transcription factor is essential for embryonic development of pulmonary vasculature.* J Biol Chem, 2005. **280**(23): p. 22278-86.
170. Krupczak-Hollis, K., et al., *The mouse Forkhead Box m1 transcription factor is essential for hepatoblast mitosis and development of intrahepatic bile ducts and vessels during liver morphogenesis.* Dev Biol, 2004. **276**(1): p. 74-88.

171. Kalin, T.V., et al., *Increased levels of the FoxM1 transcription factor accelerate development and progression of prostate carcinomas in both TRAMP and LADY transgenic mice.* Cancer Res, 2006. **66**(3): p. 1712-20.
172. Nehls, M., et al., *Two genetically separable steps in the differentiation of thymic epithelium.* Science, 1996. **272**(5263): p. 886-9.
173. Frank, J., et al., *Exposing the human nude phenotype.* Nature, 1999. **398**(6727): p. 473-4.
174. Adriani, M., et al., *Ancestral founder mutation of the nude (FOXN1) gene in congenital severe combined immunodeficiency associated with alopecia in southern Italy population.* Ann Hum Genet, 2004. **68**(Pt 3): p. 265-8.
175. Chang, J.T., et al., *Identification of differentially expressed genes in oral squamous cell carcinoma (OSCC): overexpression of NPM, CDK1 and NDRG1 and underexpression of CHES1.* Int J Cancer, 2005. **114**(6): p. 942-9.
176. Li, S., et al., *Foxn4 controls the genesis of amacrine and horizontal cells by retinal progenitors.* Neuron, 2004. **43**(6): p. 795-807.
177. Furuyama, T., et al., *Abnormal angiogenesis in Foxo1 (Fkhr)-deficient mice.* J Biol Chem, 2004. **279**(33): p. 34741-9.
178. Nakae, J., et al., *Regulation of insulin action and pancreatic beta-cell function by mutated alleles of the gene encoding forkhead transcription factor Foxo1.* Nat Genet, 2002. **32**(2): p. 245-53.
179. Hosaka, T., et al., *Disruption of forkhead transcription factor (FOXO) family members in mice reveals their functional diversification.* Proc Natl Acad Sci U S A, 2004. **101**(9): p. 2975-80.
180. Hu, H., et al., *Foxp1 is an essential transcriptional regulator of B cell development.* Nat Immunol, 2006. **7**(8): p. 819-26.
181. Shu, W., et al., *Foxp2 and Foxp1 cooperatively regulate lung and esophagus development.* Development, 2007. **134**(10): p. 1991-2000.
182. Schubert, L.A., et al., *Scurfin (FOXP3) acts as a repressor of transcription and regulates T cell activation.* J Biol Chem, 2001. **276**(40): p. 37672-9.
183. Torgerson, T.R., et al., *Severe food allergy as a variant of IPEX syndrome caused by a deletion in a noncoding region of the FOXP3 gene.* Gastroenterology, 2007. **132**(5): p. 1705-17.
184. Hong, H.K., et al., *The winged helix/forkhead transcription factor Foxq1 regulates differentiation of hair in satin mice.* Genesis, 2001. **29**(4): p. 163-71.
185. Mazet, F., C.T. Amemiya, and S.M. Shimeld, *An ancient Fox gene cluster in bilaterian animals.* Curr Biol, 2006. **16**(9): p. R314-6.
186. Mazet, F., G.N. Luke, and S.M. Shimeld, *The amphioxus FoxQ1 gene is expressed in the developing endostyle.* Gene Expr Patterns, 2005. **5**(3): p. 313-5.
187. Hunt, P. and R. Krumlauf, *Hox codes and positional specification in vertebrate embryonic axes.* Annu Rev Cell Biol, 1992. **8**: p. 227-56.
188. Aitola, M., et al., *Forkhead transcription factor FoxF2 is expressed in mesodermal tissues involved in epithelio-mesenchymal interactions.* Dev Dyn, 2000. **218**(1): p. 136-49.

189. Aranda, J., et al., *Prolactins are natural inhibitors of angiogenesis in the retina.* Invest Ophthalmol Vis Sci, 2005. **46**(8): p. 2947-53.
190. Ittner, L.M., et al., *Compound developmental eye disorders following inactivation of TGFbeta signaling in neural-crest stem cells.* J Biol, 2005. **4**(3): p. 11.
191. Ylikarppa, R., et al., *Lack of type XVIII collagen results in anterior ocular defects.* Faseb J, 2003. **17**(15): p. 2257-9.
192. Hopfer, U., et al., *Targeted disruption of Col8a1 and Col8a2 genes in mice leads to anterior segment abnormalities in the eye.* Faseb J, 2005. **19**(10): p. 1232-44.
193. Gould, D.B., R.S. Smith, and S.W. John, *Anterior segment development relevant to glaucoma.* Int J Dev Biol, 2004. **48**(8-9): p. 1015-29.
194. Thomas, S.A., A.M. Matsumoto, and R.D. Palmiter, *Noradrenaline is essential for mouse fetal development.* Nature, 1995. **374**(6523): p. 643-6.
195. Davies, A.F., et al., *An interstitial deletion of 6p24-p25 proximal to the FKHL7 locus and including AP-2alpha that affects anterior eye chamber development.* J Med Genet, 1999. **36**(9): p. 708-10.
196. Kleinjan, D.A. and V. van Heyningen, *Long-range control of gene expression: emerging mechanisms and disruption in disease.* Am J Hum Genet, 2005. **76**(1): p. 8-32.
197. Drance, S.M., *A comparison of the effects of betaxolol, timolol, and pilocarpine on visual function in patients with open-angle glaucoma.* J Glaucoma, 1998. **7**(4): p. 247-52.
198. Beebe, D.C. and J.M. Coats, *The lens organizes the anterior segment: specification of neural crest cell differentiation in the avian eye.* Dev Biol, 2000. **220**(2): p. 424-31.
199. Grimm, C., et al., *Aphakia (ak), a mouse mutation affecting early eye development: fine mapping, consideration of candidate genes and altered Pax6 and Six3 gene expression pattern.* Dev Genet, 1998. **23**(4): p. 299-316.
200. Shields, M.B., et al., *Axenfeld-Rieger syndrome. A spectrum of developmental disorders.* Surv Ophthalmol, 1985. **29**(6): p. 387-409.
201. Coghill, E.L., et al., *A gene-driven approach to the identification of ENU mutants in the mouse.* Nat Genet, 2002. **30**(3): p. 255-6.
202. Quwailid, M.M., et al., *A gene-driven ENU-based approach to generating an allelic series in any gene.* Mamm Genome, 2004. **15**(8): p. 585-91.
203. Hellqvist, M., et al., *Differential activation of lung-specific genes by two forkhead proteins, FREAC-1 and FREAC-2.* J Biol Chem, 1996. **271**(8): p. 4482-90.
204. Matthias, P., et al., *Eukaryotic expression vectors for the analysis of mutant proteins.* Nucleic Acids Res, 1989. **17**(15): p. 6418.
205. Pierrou, S., et al., *Cloning and characterization of seven human forkhead proteins: binding site specificity and DNA bending.* Embo J, 1994. **13**(20): p. 5002-12.
206. Mullis, K.B. and F.A. Faloona, *Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction.* Methods Enzymol, 1987. **155**: p. 335-50.

207. Osoegawa, K., et al., *Bacterial artificial chromosome libraries for mouse sequencing and functional analysis.* Genome Res, 2000. **10**(1): p. 116-28.
208. Orita, M., et al., *Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms.* Proc Natl Acad Sci U S A, 1989. **86**(8): p. 2766-70.
209. Liu, Q. and S.S. Sommer, *The SSCP phenomenon: addition of HEPES buffer dramatically affects electrophoretic mobility.* Biotechniques, 1998. **25**(1): p. 50-2, 54, 56.
210. Sanger, F., S. Nicklen, and A.R. Coulson, *DNA sequencing with chain-terminating inhibitors.* Proc Natl Acad Sci U S A, 1977. **74**(12): p. 5463-7.
211. Nakagata, N. and T. Takeshima, *Cryopreservation of mouse spermatozoa from inbred and F1 hybrid strains.* Jikken Dobutsu, 1993. **42**(3): p. 317-20.
212. Thornton, C.E., S.D. Brown, and P.H. Glenister, *Large numbers of mice established by in vitro fertilization with cryopreserved spermatozoa: implications and applications for genetic resource banks, mutagenesis screens, and mouse backcrosses.* Mamm Genome, 1999. **10**(10): p. 987-92.
213. Hough, T.A., et al., *Novel phenotypes identified by plasma biochemical screening in the mouse.* Mamm Genome, 2002. **13**(10): p. 595-602.
214. Waisfisz, Q., et al., *The Fanconi anemia group E gene, FANCE, maps to chromosome 6p.* Am J Hum Genet, 1999. **64**(5): p. 1400-5.
215. Kelly, A.L., et al., *Hereditary juvenile haemochromatosis: a genetically heterogeneous life-threatening iron-storage disease.* Qjm, 1998. **91**(9): p. 607-18.
216. Brzustowicz, L.M., et al., *Use of a quantitative trait to map a locus associated with severity of positive symptoms in familial schizophrenia to chromosome 6p.* Am J Hum Genet, 1997. **61**(6): p. 1388-96.
217. Gul, A., et al., *Linkage mapping of a novel susceptibility locus for Behcet's disease to chromosome 6p22-23.* Arthritis Rheum, 2001. **44**(11): p. 2693-6.
218. Chatterjee, A., et al., *Mapping the sites of putative tumor suppressor genes at 6p25 and 6p21.3 in cervical carcinoma: occurrence of allelic deletions in precancerous lesions.* Cancer Res, 2001. **61**(5): p. 2119-23.
219. Kisseljov, F., et al., *Instability of chromosome 6 microsatellite repeats in human cervical tumors carrying papillomavirus sequences.* Int J Cancer, 1996. **69**(6): p. 484-7.
220. Perou, C.M., et al., *Comparative mapping in the beige-satin region of mouse chromosome 13.* Genomics, 1997. **39**(2): p. 136-46.
221. Bailey, J.A., et al., *Analysis of segmental duplications and genome assembly in the mouse.* Genome Res, 2004. **14**(5): p. 789-801.
222. Stephenson, D.A. and K.K. Lueders, *Mouse chromosome 13.* Mamm Genome, 1999. **10**(10): p. 954.
223. Marra, M.A., et al., *High throughput fingerprint analysis of large-insert clones.* Genome Res, 1997. **7**(11): p. 1072-84.
224. Petersen, A., et al., *Expanded CAG repeats in exon 1 of the Huntington's disease gene stimulate dopamine-mediated striatal neuron autophagy and degeneration.* Hum Mol Genet, 2001. **10**(12): p. 1243-54.

225. Levinson, G. and G.A. Gutman, *Slipped-strand mispairing: a major mechanism for DNA sequence evolution.* Mol Biol Evol, 1987. **4**(3): p. 203-21.
226. Wiltshire, T., et al., *Genome-wide single-nucleotide polymorphism analysis defines haplotype patterns in mouse.* Proc Natl Acad Sci U S A, 2003. **100**(6): p. 3380-5.
227. Evans EP, D.M., *Standard Normal Chromosomes*, in *Genetic Variants and Strains of the Laboratory Mouse*, S.A. Lyon MF, Editor. 1989, Oxford University Press: Oxford. p. 576-581.
228. Furey, T.S. and D. Haussler, *Integration of the cytogenetic map with the draft human genome sequence.* Hum Mol Genet, 2003. **12**(9): p. 1037-44.
229. Bickmore, W.A. and A.T. Sumner, *Mammalian chromosome banding--an expression of genome organization.* Trends Genet, 1989. **5**(5): p. 144-8.
230. Russell, L.B., *Research News: New Mutants.* Mouse News Letters, 1955. **12**: p. 47.
231. Philip-Couderc, P., et al., *Forkhead transcription factors coordinate expression of myocardial KATP channel subunits and energy metabolism.* Circ Res, 2008. **102**(2): p. e20-35.
232. Ormestad, M., J. Astorga, and P. Carlsson, *Differences in the embryonic expression patterns of mouse Foxf1 and -2 match their distinct mutant phenotypes.* Dev Dyn, 2004. **229**(2): p. 328-33.
233. Michaud, E.J., et al., *Efficient gene-driven germ-line point mutagenesis of C57BL/6J mice.* BMC Genomics, 2005. **6**: p. 164.
234. Cox, R.D., et al., *Contrasting effects of ENU induced embryonic lethal mutations of the quaking gene.* Genomics, 1999. **57**(3): p. 333-41.
235. Austin, C.P., et al., *The knockout mouse project.* Nat Genet, 2004. **36**(9): p. 921-4.
236. Auwerx, J., et al., *The European dimension for the mouse genome mutagenesis program.* Nat Genet, 2004. **36**(9): p. 925-7.
237. Chen, Y., et al., *Genotype-based screen for ENU-induced mutations in mouse embryonic stem cells.* Nat Genet, 2000. **24**(3): p. 314-7.
238. Wienholds, E., et al., *Efficient target-selected mutagenesis in zebrafish.* Genome Res, 2003. **13**(12): p. 2700-7.
239. Schmitt, T.J., M.L. Robinson, and J. Doyle, *Single Nucleotide Polymorphism (SNP), Insertion & Deletion Detection an the WAVE Nucleic Acid Fragment Analysis System*, in *Transgenomic Application Note 112*. 2000.
240. Taylor, P., K. Munson, and D. Gjerde, *Detection of Mutations and Polymorphisms on the WAVE Nucleic Acid Fragment Analysis System*, in *Transgenomic Application Note 101*. 2000.
241. Miura, N., et al., *Mouse forkhead (winged helix) gene LUN encodes a transactivator that acts in the lung.* Genomics, 1998. **50**(3): p. 346-56.
242. Hellqvist, M., et al., *The human forkhead protein FREAC-2 contains two functionally redundant activation domains and interacts with TBP and TFIIB.* J Biol Chem, 1998. **273**(36): p. 23335-43.

243. Shapiro, M.B. and P. Senapathy, *RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression.* Nucleic Acids Res, 1987. **15**(17): p. 7155-74.
244. Carstens, R.P., W.A. Fenton, and L.R. Rosenberg, *Identification of RNA splicing errors resulting in human ornithine transcarbamylase deficiency.* Am J Hum Genet, 1991. **48**(6): p. 1105-14.
245. Robberson, B.L., G.J. Cote, and S.M. Berget, *Exon definition may facilitate splice site selection in RNAs with multiple exons.* Mol Cell Biol, 1990. **10**(1): p. 84-94.
246. Liu, J. and B. Rost, *NORSp: Predictions of long regions without regular secondary structure.* Nucleic Acids Res, 2003. **31**(13): p. 3833-5.
247. Jin, C., et al., *Dynamic DNA contacts observed in the NMR structure of winged helix protein-DNA complex.* J Mol Biol, 1999. **289**(4): p. 683-90.
248. Whelan, S. and N. Goldman, *A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach.* Mol Biol Evol, 2001. **18**(5): p. 691-9.
249. Kamprath, K. and C.T. Wotjak, *Nonassociative learning processes determine expression and extinction of conditioned fear in mice.* Learn Mem, 2004. **11**(6): p. 770-86.
250. Crawley, J. and F.K. Goodwin, *Preliminary report of a simple animal behavior model for the anxiolytic effects of benzodiazepines.* Pharmacol Biochem Behav, 1980. **13**(2): p. 167-70.
251. Davis, M., *Neural systems involved in fear and anxiety measured with fear-potentiated startle.* Am Psychol, 2006. **61**(8): p. 741-56.
252. Grillon, C., *Startle reactivity and anxiety disorders: aversive conditioning, context, and neurobiology.* Biol Psychiatry, 2002. **52**(10): p. 958-75.
253. Bowers, D., et al., *Startling facts about emotion in Parkinson's disease: blunted reactivity to aversive stimuli.* Brain, 2006. **129**(Pt 12): p. 3356-65.
254. Plappert, C.F. and P.K. Pilz, *Difference in anxiety and sensitization of the acoustic startle response between the two inbred mouse strains BALB/cAN and DBA/2N.* Genes Brain Behav, 2002. **1**(3): p. 178-86.
255. Willott, J.F. and K.R. Johnson, *Startle and Prepulse Inhibition.* 2002.
256. Braff, D.L., M.A. Geyer, and N.R. Swerdlow, *Human studies of prepulse inhibition of startle: normal subjects, patient groups, and pharmacological studies.* Psychopharmacology (Berl), 2001. **156**(2-3): p. 234-58.
257. Schwarzkopf, S.B., et al., *Test-retest reliability of prepulse inhibition of the acoustic startle response.* Biol Psychiatry, 1993. **34**(12): p. 896-900.
258. Plappert, C.F., et al., *Experience increases the prepulse inhibition of the acoustic startle response in mice.* Behav Neurosci, 2006. **120**(1): p. 16-23.
259. *Healthchex UA8 Urinalysis strips*, in *pack documentation.*
260. Brown, L.J., et al., *Normal thyroid thermogenesis but reduced viability and adiposity in mice lacking the mitochondrial glycerol phosphate dehydrogenase.* J Biol Chem, 2002. **277**(36): p. 32892-8.
261. Sherwin, R.S., R.G. Hendler, and P. Felig, *Effect of diabetes mellitus and insulin on the turnover and metabolic response to ketones in man.* Diabetes, 1976. **25**(9): p. 776-84.

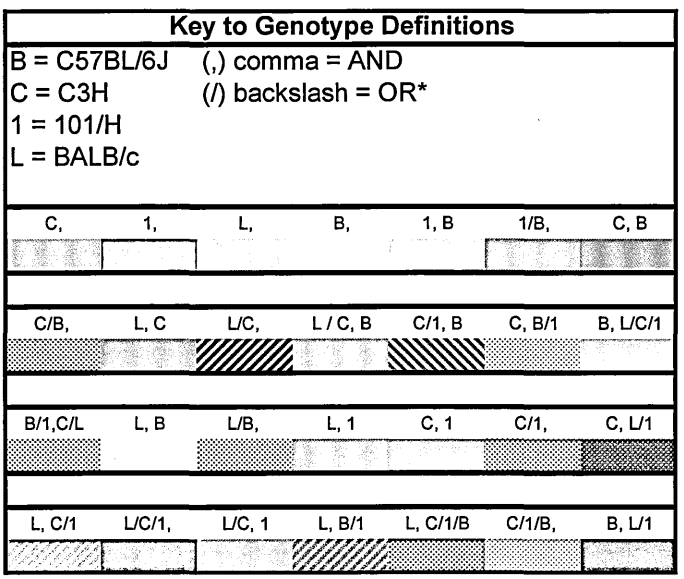
262. Pittler, S.J. and W. Baehr, *Identification of a nonsense mutation in the rod photoreceptor cGMP phosphodiesterase beta-subunit gene of the rd mouse.* Proc Natl Acad Sci U S A, 1991. **88**(19): p. 8322-6.
263. Thaung, C., et al., *Novel ENU-induced eye mutations in the mouse: models for human eye disease.* Hum Mol Genet, 2002. **11**(7): p. 755-67.
264. John, S.W., et al., *Intraocular pressure in inbred mouse strains.* Invest Ophthalmol Vis Sci, 1997. **38**(1): p. 249-53.
265. Rose, R.C., *Ascorbic acid transport in mammalian kidney.* Am J Physiol, 1986. **250**(4 Pt 2): p. F627-32.
266. Sellers, A.L., et al., *Filtration and reabsorption of protein by the kidney.* J Exp Med, 1954. **100**(1): p. 1-10.
267. Chang, B., et al., *Haploinsufficient Bmp4 ocular phenotypes include anterior segment dysgenesis with elevated intraocular pressure.* BMC Genet, 2001. **2**: p. 18.
268. Saika, S., et al., *TGFbeta2 in corneal morphogenesis during mouse embryonic development.* Dev Biol, 2001. **240**(2): p. 419-32.
269. Berry, F.B., R.A. Saleem, and M.A. Walter, *FOXC1 transcriptional regulation is mediated by N- and C-terminal activation domains and contains a phosphorylated transcriptional inhibitory domain.* J Biol Chem, 2002. **277**(12): p. 10292-7.
270. Vernes, S.C., et al., *Functional genetic analysis of mutations implicated in a human speech and language disorder.* Hum Mol Genet, 2006. **15**(21): p. 3154-67.
271. Saleem, R.A., et al., *Structural and functional analyses of disease-causing missense mutations in the forkhead domain of FOXC1.* Hum Mol Genet, 2003. **12**(22): p. 2993-3005.
272. Garner, M.M. and A. Revzin, *A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system.* Nucleic Acids Res, 1981. **9**(13): p. 3047-60.
273. Spiess, A.N., N. Mueller, and R. Ivell, *Trehalose is a potent PCR enhancer: lowering of DNA melting temperature and thermal stabilization of taq polymerase by the disaccharide trehalose.* Clin Chem, 2004. **50**(7): p. 1256-9.
274. Carninci, P., et al., *Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA.* Proc Natl Acad Sci U S A, 1998. **95**(2): p. 520-4.
275. Kain, S.R. and S. Ganguly, *Overview of genetic reporter systems.* Curr Protoc Mol Biol, 2001. **Chapter 9**: p. Unit9 6.
276. Ross, M.J., et al., *NF-kappaB regulates Fas-mediated apoptosis in HIV-associated nephropathy.* J Am Soc Nephrol, 2005. **16**(8): p. 2403-11.
277. Latham, J.P., et al., *Prostate-specific antigen promoter/enhancer driven gene therapy for prostate cancer: construction and testing of a tissue-specific adenovirus vector.* Cancer Res, 2000. **60**(2): p. 334-41.

278. Gould, D.B. and S.W. John, *Anterior segment dysgenesis and the developmental glaucomas are complex traits.* Hum Mol Genet, 2002. **11**(10): p. 1185-93.
279. Koolen, D.A., et al., *Partial iris hypoplasia in a patient with an interstitial subtelomeric 6p deletion not including the forkhead transcription factor gene FOXC1.* Eur J Hum Genet, 2005. **13**(11): p. 1169-71.
280. Alonso, A., et al., *Inhibition of T cell antigen receptor signaling by VHR-related MKPX (VHX), a new dual specificity phosphatase related to VH1 related (VHR).* J Biol Chem, 2002. **277**(7): p. 5524-8.
281. Hsiao, F.C., et al., *Eyes absent mediates cross-talk between retinal determination genes and the receptor tyrosine kinase signaling pathway.* Dev Cell, 2001. **1**(1): p. 51-61.
282. Chauhan, B.K., et al., *A comparative cDNA microarray analysis reveals a spectrum of genes regulated by Pax6 in mouse lens.* Genes Cells, 2002. **7**(12): p. 1267-83.
283. Del Bene, F., K. Tessmar-Raible, and J. Wittbrodt, *Direct interaction of geminin and Six3 in eye development.* Nature, 2004. **427**(6976): p. 745-9.
284. Chen, Q., et al., *Distinct capacities of individual E2Fs to induce cell cycle re-entry in postmitotic lens fiber cells of transgenic mice.* Dev Neurosci, 2004. **26**(5-6): p. 435-45.
285. Schilham, M.W., et al., *Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4.* Nature, 1996. **380**(6576): p. 711-4.
286. Critcher, R., et al., *Assignment of Sox4 to mouse chromosome 13 bands A3-A5 by fluorescence in situ hybridization; refinement of the human SOX4 location to 6p22.3 and of SOX20 to chromosome 17p12.3.* Cytogenet Cell Genet, 1998. **81**(3-4): p. 294-5.
287. van der Knaap, M.S., et al., *Cerebral white matter abnormalities in 6p25 deletion syndrome.* AJNR Am J Neuroradiol, 2006. **27**(3): p. 586-8.
288. Giebisch, G., *Renal potassium transport: mechanisms and regulation.* Am J Physiol, 1998. **274**(5 Pt 2): p. F817-33.
289. Sharma, N., et al., *Familial hyperinsulinism and pancreatic beta-cell ATP-sensitive potassium channels.* Kidney Int, 2000. **57**(3): p. 803-8.
290. DeFronzo, R.A., et al., *The effect of insulin on renal handling of sodium, potassium, calcium, and phosphate in man.* J Clin Invest, 1975. **55**(4): p. 845-55.
291. Willcutt, E.G., et al., *Quantitative trait locus for reading disability on chromosome 6p is pleiotropic for attention-deficit/hyperactivity disorder.* Am J Med Genet, 2002. **114**(3): p. 260-8.
292. Nash, M.W., et al., *Genome-wide linkage analysis of a composite index of neuroticism and mood-related scales in extreme selected sibships.* Hum Mol Genet, 2004. **13**(19): p. 2173-82.
293. Seong, E., A.F. Seasholtz, and M. Burmeister, *Mouse models for psychiatric disorders.* Trends Genet, 2002. **18**(12): p. 643-50.
294. Bland, J.M. and D.G. Altman, *Multiple significance tests: the Bonferroni method.* Bmj, 1995. **310**(6973): p. 170.

295. Davis, M., et al., *Fear-potentiated startle: a neural and pharmacological analysis.* Behav Brain Res, 1993. **58**(1-2): p. 175-98.
296. Ludewig, S., et al., *Information-processing deficits and cognitive dysfunction in panic disorder.* J Psychiatry Neurosci, 2005. **30**(1): p. 37-43.
297. Grillon, C., et al., *Baseline startle amplitude and prepulse inhibition in Vietnam veterans with posttraumatic stress disorder.* Psychiatry Res, 1996. **64**(3): p. 169-78.
298. Duncan, G.E., et al., *Deficits in sensorimotor gating and tests of social behavior in a genetic model of reduced NMDA receptor function.* Behav Brain Res, 2004. **153**(2): p. 507-19.
299. Jaubert, P.J., et al., *Complex, multimodal behavioral profile of the Homer1 knockout mouse.* Genes Brain Behav, 2007. **6**(2): p. 141-54.
300. Sallinen, J., et al., *Adrenergic alpha2C-receptors modulate the acoustic startle reflex, prepulse inhibition, and aggression in mice.* J Neurosci, 1998. **18**(8): p. 3035-42.
301. Kumari, V., et al., *Association between violent behaviour and impaired prepulse inhibition of the startle response in antisocial personality disorder and schizophrenia.* Behav Brain Res, 2005. **158**(1): p. 159-66.
302. Koch, M., *The neurobiology of startle.* Prog Neurobiol, 1999. **59**(2): p. 107-28.
303. Benjamini, Y., et al., *Controlling the false discovery rate in behavior genetics research.* Behav Brain Res, 2001. **125**(1-2): p. 279-84.
304. Benjamini, Y. and W. Liu, *A Distribution-Free Multiple Test Procedure That Controls The False Discovery Rate.* 1999.
305. Benjamini, Y.H., Y., *Controlling the false discovery rate: a practical and powerful approach to multiple testing.* Journal of the Royal Statistical Society, 1995. **57**(1): p. 289-300.

Appendices

Appendices for Chapter 3
Appendix 3.1



Key to genotyping data in appendix 3.1. *NB: (,) comma equals "AND" and (/) backslash represents other potential alleles of the same size. This would usually equate to "OR", but when used for an individual with only one allele size (no strains after the comma), the backslash equates to "AND/OR".

Appendix 3.1

Genotyping Data For All Individuals

Mouse ID	No. Individual	d13Mit115	d13Mit17	DNR059	DNR123	d13Mit15	d13Mit133	DNR48	DNR76	DNR101	TetNR5	TetNR10	DNR037	DNR24	DNR003	d13Mit136	DNR16	d13Mit198
D36/1409.01	2	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
D36/1447.01	4	1,B	1,B	1,B							1,B	1,B	1,B	1,B	1,B	1,B	1,B	
D36/1463.01	1	1,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		1,B
D36/1463.02	1	C,B	C,B		B,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
D36/1463.03	3	C,B	C,B		C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
D36/1463.04	1	C,B	C,B		1,B	C/1,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
D36/1525.01	2	1,B	1,B		1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B		1,B		1,B
D36/1525.02	2	1,B	1,B		1,B		1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B		1,B		1,B
D36/1525.03	1	1,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
D36/1525.04	1	1,B	B,		B,		B,	B,	B,	B,	B,	B,	B,	B,		B,		1,B
D36/1558.01	6	1,B	1,B		1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B		1,B		1,B
D36/1558.02	1	1,B	B,		B,		B,	B,	B,	B,	B,	B,	B,	B,		B,		1,B
D36/1558.03	1	C,B	C,B		C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
D36/1560.01	1	1,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
D36/1560.02	2	1,B	B,		B,		B,	B,	B,	B,	B,	B,	B,	B,		B,		1,B
D36/1560.03	1	C,B	C,B		C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
D36/1561.01	1	1,B	C,B		B,		B,	B,	B,	C,B	C,B	B,	B,	B,		B,		B,
D36/1561.01	1	C,B	C,B		B,		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
D36/1562.01	1	C,B	C,B		C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M1073.01	1	B,	B,		B,		B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M1073.02	3	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M1073.03	4	B,	B,	B,							B,	B,	B,	B,	B,	B,	B,	
M1073.04	2	1,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M1073.05	1	1,B	B,		C,B		B,	B,	B,	B,	B,	B,	B,	B,		B,		1,B
M1073.06	4	1,B	B,		B,		B,	B,	B,	B,	B,	B,	B,	B,		B,		1,B
M1073.07	1	1,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M1073.08	5	1,B	B,	B,							B,	B,	B,	B,	B,	B,	B,	
M1073.09	7	C,B	C,B		B,		B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M1073.10	2	C,B	C,B		C,B	C,B	C,B	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M1073.11	1	C,B	C,B	C,B							C,B	C,B	C,B	B,	B,	B,	B,	
M1073.12	1	C,1	C,		C,	C,	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,B/1
M1073.13	2	C,1	C,		C,	C,	C,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M1073.14	1	C,1	C,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M1073.15	3	C,1	C,	C,							C,	C,	C,	C,	C,	C,	C,	
M1073.16	6	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
M1073.17	1	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M1185.01	6	1,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M1185.02	1	1,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		1,B
M1185.03	4	1,B	B,		B,		B,	B,	B,	B,	B,	B,	B,	B,		B,		1,B
M1185.04	2	1,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M1185.05	20	1,B	B,	B,							B,	B,	B,	B,	B,	B,	B,	
M1185.06	1	1,B	B,	B,							C,	C,	C,	C,	C,	C,	C,	
M1185.07	1	B,	B,		B,		B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M1185.08	2	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		
M1185.09	11	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M1185.10	5	B,	B,		B,		B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M1185.11	15	B,	B,	B,							B,	B,	B,	B,	B,	B,	B,	
M1185.12	1	B,	B,		B,	B,	B,	B,	B,	B,	B,	C,B	C,B	C,B		C,B		C,B
M1185.13	1	B,	B,	B,							B,	C,B	C,B	C,B	C,B	C,B	C,	
M1185.14	2	B,	B,	B,							B,	C,B	C,B	C,B	C,B	C,B	C,B	
M1185.15	1	B,	B,		B,	B,	B,	B,	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M1185.16	1	B,	B,		B,	B,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		
M1185.17	1	C,1	C,	C,1							B,	B,	B,	B,	B,	B,	B,	
M1185.18	1	C,1	C,	C,							C,	C,	B,	B,	B,	B,	B,	
M1185.19	4	C,1	C,		C,	C,	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,B/1
M1185.20	1	C,1	C,		C,	C,	C,1	C,	C,	C,	C,	C,	C,	C,		C,		C,B/1
M1185.21	3	C,1	C,		C,		C,	C,	C,	C,	C,	C,	C,	C,		C,		C,1
M1185.22	1	C,1	C,		C,	C,	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,1
M1185.23	1	C,1	C,		C,	C,	C,	C,	C,	C,	C,	C,	C,	C,		C,		
M1185.24	5	C,1	C,	C,							C,	C,	C,	C,	C,	C,	C,	
M1185.25	3	C,B	C,B		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M1185.26	1	C,B	L/C,B		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M1185.27	2	C,B	C,B		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		
M1185.28	1	C,B	C,B		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M1185.29	2	C,B	C,B		B,		B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M1185.30	4	C,B	C,B	C,B							B,	B,	B,	B,	B,	B,	B,	
M1185.31	1	C,B	C,B		B,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		
M1185.32	8	C,B	C,B		C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M1185.33	14	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M1185.34	2	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		
M1185.35	1	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B			C,B		C,B
M1185.36	13	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
M1239.01	2	1,	1,		1,		1,	1,	1,	1,B	1,	1,	1,	1,		1,		1,

Mouse ID	No. Individuals	d13Mit115	d13Mit17	DNR059	DNR123	d13Mit15	d13Mit133	DNR48	DNR76	DNR101	TetNR5	TetNR10	DNR037	DNR24	DNR003	d13Mit136	DNR16	d13Mit198
M1239.02	2	1,	1,	1,	,	,	,	,	,	,	1,	1,	1,	1,	1,	1,	1,	,
M1239.03	1	1,	1,	1,	,	,	,	,	,	,	1,	1,	1,	,	,	1,	1,	,
M1239.04	3	1,B	1,B	,	1,B	,	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	,	1,B	,	1,B
M1239.05	7	1,B	1,B	1,B	,	,	,	,	,	,	1,B	1,B	1,B	1,B	1,B	1,B	1,B	,
M1239.06	1	B,	1,B	1,B	,	,	,	,	,	,	1,B	1,B	1,B	1,B	1,B	1,B	1,B	,
M1239.07	1	1,B	1,B	,	1,B	1,B	1,B	1,B	1,B	B,	1,B	1,B	1,B	1,B	,	1,B	,	1,B
M1239.08	1	1,B	1,B	,	1,B	1,B	1,B	1,B	1,B	B,	B,	1/B,	1,B	1,B	,	1/B,	,	1/B,
M1239.09	1	1,B	1,B	,	1,B	1,B	1,B	1,B	1,B	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M1239.10	10	1,B	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M1239.11	1	1,B	,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M1239.12	2	1,B	B,	,	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	,	B,	,	B,
M1239.13	9	1,B	B,	,	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	,	B,	,	1,B
M1239.14	1	B,	1,B	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M1239.15	8	B,	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M1239.16	1	B,	B,	,	C,B	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M1239.17	1	B,	B,	,	B,	B,	B,	C,B	C,B	C,B	C,B	B,	B,	B,	,	1/B,	,	1/B,
M1239.18	8	B,	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	B,	,	B,
M1239.19	7	B,	B,	,	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	,	B,	,	B,
M1239.20	1	B,	,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	B,	,	B,
M1239.21	2	L,B	L,B	,	L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B	,	L,B	,	L,B
M1239.22	1	L,C	L,C	,	L,C	L,C	L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B	,	L,B	,	L,B
M1239.23	5	L,C	L,C	,	L,C	L,C	L,C	L,C	L,C	L,C	L,C	L,C	L,C	L,C	,	L,C	,	L,C
M1239.24	1	C,1	C,	,	C,	C,	C,	C,	C,	C,	C	B,	B,	B,	,	1/B,	,	1/B,
M1239.25	2	C,1	C,	,	C,	C,	C,	C,	C,	C,B	C,	B,	B,	B,	,	1/B,	,	1/B,
M1239.26	1	C,1	C,	,	,	C,	C,	C,	C,	C,B	C,	B,	B,	B,	,	1/B,	,	1/B,
M1239.27	1?	C,1	C,	,	C,	C,	C,	C,	C,	B,	C,	211.35	B,	B,	,	1/B,	,	1/B,
M1239.28	2	C,1	C,	,	C,	C,	C,	C,	C,	B,	C,	B	B,	B,	,	1/B,	,	1/B,
M1239.29	2	C,1	C,1	,	1,	C/1	C,1	C,1	C/1	B,	C,	1/B,	1,B	1,B	,	1/B,	,	1/B,
M1239.30	1	C,1	C,	,	,	C,	C,	C,	C,	C,	C,	C,	C,	C,	,	C,	,	B/1,L/C
M1239.31	1	C,1	C,	,	C,	C,	C,	C,	C,	C,	C,	C,	C,	C,	,	C,	,	C,1
M1239.32	1	C,1	C,	,	C,	,	C,	C,	C,	C,	C,	C,	C,	C,	,	C,	,	C,1
M1239.33	1	C,1	C,	C,	,	,	,	,	,	,	C,	C,	C,	C,	C,	C,	C,	,
M1239.34	1	C,1	C,	C,	,	,	,	,	,	,	C,1	C,	C,	C,	C,	C	C,1	,
M1239.35	1	C,B	C,B	,	C,B	C,B	C,B	C,B	B,	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M1239.36	1	C,B	C,B	,	C,B	C,B	C,B	C,B	C,B	B,	C,B	1/B,	B,	B,	,	1/B,	,	1/B,
M1239.37	2	C,B	C,B	,	C,B	C,B	C,B	C,B	C,B	B,	C,B	B,	B,	B,	,	1/B,	,	1/B,
M1239.38	1	C,B	C,	,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	B,	B,	,	1/B,	,	1/B,
M1239.39	5	C,B	C,B	,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	B,	B,	B,	,	1/B,	,	1/B,
M1239.40	1	C,B	C,B	,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	B,	,	C,B	,	C,B
M1239.41	1	C,B	C,	,	,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	,	C,B	,	C,B
M1239.42	1	C,B	C,	C,B	,	,	,	,	,	,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	,
M1239.43	11	C,B	C,B	,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	,	C,B	,	C,B
M1239.44	1	C,B	C,B	,	C,B	,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	,	C,B	,	C,B
M1239.45	13	C,B	C,B	C,B	,	,	,	,	,	,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	,
M1616.01	1	B,	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	B,	,	B,
M1616.02	1	1,B	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M1616.03	1	C,1	C,	,	C,	C,	C,	C,	C,	C,	C,	C,	C,	C,	,	C,	,	C,B/1
M1616.04	2	C,B	C,B	,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	,	C,B	,	C,B
M1616.05	1	C,B	C,B	,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	,	,	C,B	,	C,B
M1645.01	4	1,B	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M1645.02	1	B,	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M1645.03	14	B,	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	B,	,	B,
M239.01	1	1,B	1,B	,	1,B	1,B	C,1	C,1	C/1,	C/1,	1,	C,B/1	C,1	C,1	,	C,1	,	C,1
M239.02	2	1,B	1,B	,	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	,	1/B,	,	1/B,
M239.03	1	1,B	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M239.04	1	1,B	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	B/1,C/L
M239.05	1	1,B	1,B	,	1,B	C/1,B	C,B	C,1	C/1,	C/1,	C,1	C,1	C,1	C,1	,	C,1	,	C,1
M239.06	1	1,B	1,B	,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	,	C,B	,	C,B
M239.07	5	B,	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	B,	,	B,
M239.08	2	B,	B,	,	B,	B,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	,	C,B	,	C,B
M239.09	1	C,1	C,	,	C,	C,	C,	C,	C,	C,	C,	C,	C,	C,	,	C,	,	C,B/1
M239.10	2	C,B	C,B	,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	,	C,B	,	C,B
M241.01	1	1,B	B,	,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M241.02	26	1,B	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M241.03	7	B,	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M241.04	1	B,	,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M241.05	1	1,B	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	B/1,C/L
M241.06	1	1,B	B,	,	B,	L,	B,	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M241.07	7	1,B	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	B,	,	1/B,
M241.08	1	1,B	B,	,	B,	B,	B,	B,	B,	B,	B,	L/C,B	B,	B,	,	B,	,	1/B,
M241.09	1	B,	B,	,	B,	B,	C,B	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M241.10	1	B,	B	,	C,B	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	1/B,	,	1/B,
M241.11	5	1,B	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	B,	,	B,
M241.12	6	B,	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	B,	,	1/B,
M241.13	56	B,	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	B,	,	B,	,	B,
M241.14	1	1,B	B,	,	B,	B,	B,	B,	B,	B,	B,	B,	B,	///	,	///	,	B/1,C/L
M241.15	1	1,B	B,	,	B,	,	C,	C,	C,	C,	C,	C,	C,	C,	,	C,	,	C,1
M241.16	1	1,B	B,	,	B,	B,	C,	C,	C,	C,	C,	C,	C,B	C,	,	C,	,	C,B/1

Mouse ID	No. Individuals	d13Mit115	d13Mit17	DNR059	DNR123	d13Mit15	d13Mit133	DNR48	DNR76	DNR101	TetNR5	TetNR10	DNR037	DNR24	DNR003	d13Mit136	DNR16	d13Mit198
M241.17	1	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	C,B		C,B		C,B
M241.18	4	B,	B,		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M241.19	1	B,	B,		B,	B,	B,	B,	B,	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M241.20	5	B,	B,		B,	B,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M241.21	1	L,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M241.22	2	L,B	L,B		L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B		L,B		L,B
M241.23	2	L,C	L,C		L,C	L,C	L,C	L,C	L,C	L,C	L,C	L,C	L,C	L,C		L,C		L,C
M241.24	1	L,C	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		L/C,B		L/C,B
M241.25	2	C,	C,		C,	C,	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,
M241.26	3	C,1	C,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M241.27	6	C,1	C,		C,	C,	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,B/1
M241.28	5	C,1	C,		C,	C,	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,1
M241.29	1	C,B	L/C,B		C,B	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M241.30	1	C,B	C,B		C,B	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M241.31	8	C,B	C,B		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M241.32	1	C,B	C,B		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M241.33	1	C,B	C,B		C,B	C,B	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,
M241.34	1	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	B,		1/B,		1/B,
M241.35	1	C,B	C,		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M241.36	1	C,B	C,			C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B			C,B		C,B
M241.37	2	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		1/B,
M241.38	1	C,B	C,B		C,B	L C/1/B	L/C,B	L/C,B	L/C,B	L/C/1,B	C,B	C,B	C,B	C,B		C,B		C,B
M241.39	1	C,B	C,B		B,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B			C,B		C,B
M241.40	51	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M241.41	1	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B			C,B		C,B
M241.42	1	C,B	L/C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M369.01	2	1,B	1,B		1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B		1,B		1,B
M369.02	1	1,B	1,B		1,	C/1/B,	C,1	C,1	C/1	C/1,B	C,1	1/B,	1,B	1,B		1/B,		1/B,
M369.03	1	1,B	1,B		1,	C,1,B	C,1	C,1	C/1	C/1	C,1	C,1	C,1	C,1		C,1		C,1
M369.04	1	1,B	1,B		1,	C/1,	C,1	C,1	C/1	C/1	C,1	C,1	C,1	C,1		C,1		C,1
M369.05	1	1,B		B,								B,	B,	B,	B,	B,	B,	
M369.06	2	1,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M369.07	1	1,B	B,	B,							B,	B,				B,		
M369.08	7	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M369.09	1	B,		B,								B,	B,	B,	B,	B,	B,	
M369.10	1	B,		B,								B,				B,		
M369.11	1	B,	B,	B,							B,	B,	B,			B,	B,	
M369.12	1	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		L/C,B
M369.13	1	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	L/C,B		L/C,B		L/C,B
M369.14	1	B,	B,		C,B	C,B	C,B	C,B	C,B	C,B	B,	C,B	C,B	C,B		C,B		C,B
M369.15	1	B,	B,		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M369.16	1	B,	B,	B,							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
M369.17	1	C,1	C,		C,	C,	C,	C,	C,	C,B	B,	B,	B,	B,		1/B,		1/B,
M369.18	2	C,1	C,		C,	C,	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,B/1
M369.19	1	C,1	C,	C,							C,	C,	C,	C,	C,	C,	C,	
M369.20	1	C,1	C,1		C,1	C,1	C,1	C,1	C,1	C,1	C,1	C,1	C,1	C,1		C,1		C,1
M369.21	1	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	B,		1/B,		1/B,
M369.22	10	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M369.23	2	C,B	C,B	C,B							C,B	C,B	C,B			C,B	C,B	
M369.24	1	C,B	C,B	C,B							C,B	C,B				C,B		
M369.25	1	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
M412.01	1	1,B	1,B		1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B		1,B		1,B
M412.02	1	1,B	1,B		1,B	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M412.03	3	1,B			B,		L/C,B	B,	B,	B,			B,	L/C,B		1/B,		1/B,
M412.04	1	1,B	B,		B,	B,	B,	B,	B,	B,	L/B	1/B,	B,	B,		1/B,		1/B,
M412.05	25	1,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M412.06	3	1,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		
M412.07	1	1,B			C,		L/C,1	C/1	L/C/1,	L/C/1,			C,	C/1		C/1		B/1,C/1
M412.08	2	1,B	B,		C,	B,	C,	C,	C,	C,	C,	C,	C,	C,		C,		B/1,C/1
M412.09	1	1,B	B,		C,	B,	C,	C,1	C,	C,	C,	C,	C,	C,		C,		C,B/1
M412.10	1	1,B	B,		C,	B,	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,1
M412.11	2	1,B	B,		C,	B,	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,B/1
M412.12	1	1,B	1,B		C,	1,B	1,B	1,B	1,B	1,B	1,B	C,1	C,1	C,1		C,1		C,1
M412.13	1	1,B	1,B		1,B	1,B	1,B	1,B	1,B	1,B	1,B	C,1	C,1	C,1		C,1		C,1
M412.14	2	B,			B,		L/C,B	B,	B,	B,			B,	L/C,B		1/B,		1/B,
M412.15	1	B,			B,		L/C,B	B,	B,	B,			B,	B,		1/B,		1/B,
M412.16	1	B,	B,		B,	B,	L/C,B	B,	B,	B/1,C/1	L/B	B,	B,	L/C,B		1/B,		1/B,
M412.17	3	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M412.18	26	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M412.19	10	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		
M412.20	1		B,			B,		B,	B,	B,	B,	B,	B,	B,		B,		B,
M412.21	1	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	L/C,B		L/C,B		
M412.22	1	B,	B,		B,	1,B	1,B	1,B	1,B	1,B	1,B	C,1	C,1	C,1		C,1		C,1
M412.23	1	B,	B,		B,	B,	B,	B,	B,	B,L/C/1	L/B,	B/1,C/1	C,B	L/C,B		B/1,C/1		B/1,C/1
M412.24	1	C,	1,B		1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B		1,B		1/B,
M412.25	1	C	C/1		C,L/1	L,C/1	C/1	C/1	C/1	C/1	L/B,	B,	B,	B,		1/B,		1/B,
M412.26	1	C,1	C,		C,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M412.27	3	C,1			C,		L/C,B	B,	B,	B,			B,	L/C,B		1/B,		1/B,

Mouse ID	No. Individuals	d13Mit115	d13Mit17	DNR059	DNR123	d13Mit15	d13Mit133	DNR48	DNR76	DNR101	TetNR5	TetNR10	DNR037	DNR24	DNR003	d13Mit136	DNR16	d13Mit198
M412.28	1	C,1			C,		L/C,	L/C,	L/C/1,	L/C/1,			C,	L/C,		L/C,		B/1,C/1,
M412.29	2	C,1	C,		C,	L,	L/C,	L/C,	L/C,	L/C,	C,	C,	C,	C,		C,		
M412.30	3	C,1	C,		C,	C,	C,	C,	C,	C,	C,	C,	C,	C,		C,		
M412.31	1	C,1	C,		C,	B,	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,1
M412.32	2	C,1	C,		C,	B,	C,	C,	C,	C,	C,	C,	C,	C,		C,		B/1,L/C
M412.33	3	C,1	C,		C,	B,	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,B/1
M412.34	1	C,1	C,			B,	C,	C,	C,	C,	C,	C,	C,			C,		C,B/1
M412.35	4	C,B	C,B		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M412.36	4	C,B	C,B		C,B	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M412.37	1	C,B	C,B		C,B	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M412.38	1	C,B			C,B		L/C,B	B	B,	B,			B,	L/C,B		1/B,		1/B,
M412.39	5	C,B			C,B		L/C,B	L/C,B	B,L/C/1	L/C/1,E			C,B	L/C,B		B/1,C/1,		B/1,C/1,
M412.40	1	C,B	C,B		C,B	B,	C,B	C,B	C,B	C,	C,	C,	C,	C,		C,B/1		
M412.41	1	B,	B,		B,	B,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		
M412.42	2	B,	B,		B,	B,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M412.43	1	B,	B,		B,	B,	B,	B,	B,	B,	B,	C,B	C,B	C,B		C,B		C,B
M412.44	14	B,	B,		C,B	B,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		
M412.45	2	C,B	C,B		B,	B,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M412.46	1	C,B	C,B		C,B	B	C,B	C,B	C,B	C,B	B,	C,B	C,B	C,B		C,B		C,B
M412.47	21	C,B	C,B		C,B	B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M412.48	1	C,B	C,B			B,	C,B	C,B	C,B	C,B	C,B	C,B	C,B			C,B		C,B
M412.49	2	C,B	C,B		C,B	B,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		
M412.50	4	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M412.51	1	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		
M54.01	18	1,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M54.02	1	1,B	B,		B,	B,	B,	B,	B,		B,	B,	B,	B,		1/B,		1/B,
M54.03	3	1,B			B,		L/C,B	B,	B,	B,			B,	B,		1/B,		1/B,
M54.04	8	1,B			B,		L/C,B	B,	B,	B,			B,	L/C,B		1/B,		1/B,
M54.05	4	1,B			B,L/1		L/C,B	1,B	L/C/1,B	L/C/1,E			B,L/1	L/C,B		1/B,		1/B,
M54.06	1	1,B			B,L/1		L/C,B	1,B	L/C/1,B	L/C/1,E			B,L/1	1,B		1/B,		1/B,
M54.07	1	1,B	B,		B,	B,	B	B,	B,	B,	B,	B,	B,	C,		C,		C,1
M54.08	5	B,			B,		L/C,B	B,	B,	B,			B,	L/C,B		1/B,		1/B,
M54.09	2	B,			B,		L/C,B	B,	B,	B,			B,	B,		1/B,		1/B,
M54.10	1	B,			B,		B,	B,	B,	B,			B,	B,		1/B,		1/B,
M54.11	2	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M54.12	7	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M54.13	1	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		1/B,
M54.14	3	B,	B,		B,	B,	B	B,	B,	B,	C,B	C,B	C,B	C,B		C,B		C,B
M54.15	1	B,			B,		L/C,B	B,	B,	B,			C,B	L/C,B		B/1,C/1,		B/1,C/1,
M54.16	1	1,B			B,		L/C,	L/C,	L/C/1,	L/C/1,			C,	L/C,		L/C,		B/1,C/1,
M54.17	3	C,1			C,		L/C,	L/C,	L/C/1,	L/C/1,			C,	L/C,		L/C,		B/1,C/1,
M54.18	1	C,1			C,		L/C,	L/C,	L/C/1,	L/C/1,			C,	L/C,B		1/B,		1/B,
M54.19	1	C,1			C,		L/C,1	L/C,	L/C/1,	L/C/1,			C,	L/C,B		1/B,		1/B,
M54.20	1	C,1			C,1		L/C,B	L/C,1	L/C/1,	L/C/1,			C,B/1	L/C,1		B/1,C/1,		B/1,C/1,
M54.21	3	C,1			C,B/1		L/C,B	L/C,1	L/C/1,	L/C/1,			C,B/1	L/C,1		B/1,C/1,		B/1,C/1,
M54.22	2	C,1	C,		C,	B,	C,	C,	C,	C,	C,	C,	C,	B,		1/B,		1/B,
M54.23	1	C,1	C,		C,	B,	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,1
M54.24	1	C,1	C,		C,	B,	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,B/1
M54.25	1	C,B			C,B		L/C,B	L/C,B	B,	B,			B,	L/C,B		1/B,		1/B,
M54.26	1	C,1			B,L/1		L/C,B	L/C,B	L/C/1,B	L/C/1,B			C,B	L/C,B		B/1,C/1,		B/1,C/1,
M54.27	3	C,B			C,B		L/C,B	L/C,B	B,L/C/1	L/C/1,B			C,B	L/C,B		1/B,		1/B,
M54.28	3	B,			C,B		L/C,B	L/C,B	B,L/C/1	L/C/1,B			C,B	L/C,B		B/1,C/1,		B/1,C/1,
M54.29	10	C,B			C,B		L/C,B	L/C,B	B,L/C/1	L/C/1,B			C,B	L/C,B		B/1,C/1,		B/1,C/1,
M54.30	1	C,1			C,B/1		L/C,B	1,B	B,L/C/1	L/C/1,B			B,L/1	1,B		1/B,		1/B,
M54.31	1	C,B	C,B		C,B	C,B	C,B	B,	B,	B,	B	B,	B,	B,		B,		B,
M54.32	1	C,B	C,B		C,B	B,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	B,		1/B,		1/B,
M54.33	8	C,B	C,B		C,B	B,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M54.34	4	C,B	C,B		B,	B,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M54.35	1	C,B	C,B		C,B	C,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M54.36	3	C,B			C,B		C,B	C,B	C,B	C,B			C,B	C,B		C,B		C,B
M54.37	3	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M624.01	1	1,B	1,B		1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B		1,B		1,B
M624.02	1	1,B	1,B		1,B		1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B		1,B		1,B
M624.03	1	1,B	1,B		1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B		B/1,L/C		1,B
M624.04	1	1,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		,B/1		,B/1
M624.05	14	1,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M624.06	2	1,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		1,B
M624.07	2	1,B	B,		B,		B,	B,	B,	B,	B,	B,	B,	B,		B,		1,B
M624.08	2	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M624.09	21	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M624.10	2	B	B,		B,	B,	B,	B,	B,	B,	B	B,	B,	B,		B,		
M624.11	3	B	B,		B,		B,	B,	B,	B,	B	B,	B,	B,		B,		B,
M624.12	1	1,B	B,		B,	B,	B,	B,	B,	C/1,B	B,	C,	C,	C,		C,		C,1
M624.13	3	B,	B,		B,	B,	B,	B,	B,	B,	C,B	C,B	C,B	C,B		C,B		C,B
M624.14	2	L,B	L,B		L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B		L,B		L,B
M624.15	4	L,C	L,C		L,C	L,C	L,C	L,C	L,C	L,C	L,C	L,C	L,C	L,C		L,C		L,C
M624.16	1	C,1	C,1		C,1	C,1/B,	1,B	1,B	1,B	1,B	1,B	1,B	1,B	1,B		1/B,		1/B,
M624.17	1	C,1	C,1		C,1	C,1	C,1	C,1	C,1	C,1	C,1	C,B/1	1,B	C,B		1/B,		1/B,

Mouse ID	No. Individuals	d13Mit115	d13Mit17	DNR059	DNR123	d13Mit15	d13Mit133	DNR48	DNR76	DNR101	TetNR5	TetNR10	DNR037	DNR24	DNR003	d13Mit136	DNR16	d13Mit198
M624.18	1	C,1	C,		C,	C,	C,	C,	C,	C,B	B,	B,	B,	B,		B,		B,
M624.19	1	C,1	C,		C,	C,	C,	C,	C,	C,	C,	C,	B	B,		B,		
M624.20	1	C,1	C,1		1,	C,1	C,1	C,1	C,1	C,1	C,1	C/1,C/1	B,L/1	L/C,1		B/1,C/1		
M624.21	4	C,1	C,		C,	C,	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,B/1
M624.22	1	C,1	C,1		C,1	C,1	C,1	C,1	C,1	C,1	C,1	C,1	C,1	C,1		C,1		C,1
M624.23	1	C,1	C,1		1,		C,1	C,1	C,1	C/1,B	C,1	C,1	C,1	C,1		C,1		C,1
M624.24	1	C,1	C,		C,	C,	C,	C,	C,	C,	C,	C,	C,B	C,		C,		C,B
M624.25	2	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		B,		B,		
M624.26	1	C,B	C,B		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M624.27	1	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	B,	B,		1/B,		1/B,
M624.28	3	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	B,	B,		B,		B,
M624.29	2	C,B	C,B		C,B		C,B	C,B	C,B	C,B	C,B	C,B	B,	B,		B,		B,
M624.30	1	C,B	C,		B,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M624.31	11	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M624.32	1	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		
M624.33	3	C,B	C,B		C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M667.01	12	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M667.02	2	L,	L,		L,	L,	L,	L,	L,	L,	L,	L,	L,	L,		L,		L,
M667.03	1	B,	L,		L,	L,	L,	L,	L,	L,	L,	U/U	C,L/1	U/U		U/U		U/U
M667.04	1	L,	L,C		L,C	L,C	L,C	L,C	L,C	L,C	L,C	L,C	L,C	L,C		L,C		L,C
M667.05	1	L,B	L,		L,B	C/1,B	L/C,B	L/C,B	L/C,B	L/C,B	L,B	L,B	L,B	L,B		L,B		L,B
M667.06	3	L,B	L,		L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B		L,B		L,B
M667.07	5	L,B	L,B		L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B		L,B		L,B
M667.08	1	C,B	C,B		B,L/1	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M667.09	2	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	B,	B,		B,		B,
M667.10	3	C,B	C,		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M876.01	6	1,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M876.02	3	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
M876.03	1	B,		B,							B,	B,	B,	B,	B,	B,	B,	
M876.04	16	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M876.05	2	C,B	C,B		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
M876.06	4	C,1	C,		C,	C,	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,B/1
M876.07	1	C,	C,		C,	C,	C,	C,	C,	C,	C,	C,B	C,B	C,B		C,B		C,B
M876.08	1	C,	C,		C,	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M876.09	1	C,	C,		C,	C/1,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M876.10	4	C,B	C,		C,B	C/1,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M876.11	2	C,B	C,	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
M876.12	2	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
M876.13	1		C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
M876.14	17	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
Mud/1368c.01	3	L,1	L,	L,							L,	L,	L,	L,	L,	L,	L,	
Mud/1368c.02	4	L,B	L,B	L,B							L,B	L,B	L,B	L,B	L,B	L,B	L,B	
Mud/1368c.03	8	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mud/1368c.04	2	C,1	C,	C,							C,	C,	C,	C,	C,	C,	C,	
Mud/1528c.01	3	L,1	L,	L,							L,	L,	L,	L,	L,	L,	L,	
Mud/1528c.02	1	C,1	C,	C,							C,	C,	C,	C,	C,	C,	C,	
Mud/1528c.03	1	L,B	L,B	L,B							L,B	L,B	L,B	L,B	L,	L,B	L,B	
Mud/1528c.04	1	C,B	L/C,B	L/C,B							L,B	L,B	L,B	L,B	L,B	L,B	L,B	
Mud/1528c.05	1	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mud/1603.01	2	L,1	L,	L,							L,	L,	L,	L,	L,	L,	L,	
Mud/1603.02	1	L,1	L,	L,							L,	L,	L,			L,	L,	
Mud/1603.03	1	L,B	L,	L,B							L,B	L,B	L,B	L,B	L,B	L,B	L,B	
Mud/1603.04	5	L,B	L,B	L,B							L,B	L,B	L,B	L,B	L,B	L,B	L,B	
Mud/1603.05	1	L,B	L,B	L,B							L,B	L,B		L,B	L,B	L,B		
Mud/1603.06	1	L,B	L,B	L,B							L,B	L,B	L,B			L,B	L,B	
Mud/1603.07	1	L,B	L,B	L,B							L,B	L,B	L,B	L,B	L,	L,B	L,B	
Mud/1603.08	3	C,1	C,	C,							C,	C,	C,	C,	C,	C,	C,	
Mud/1603.09	1	C,1	C,	C,							C,	C,	C,B	C,	C,	C,	C,	
Mud/1603.10	2	C,B	C,	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mud/1603.11		C,B	C,	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mud/1603.12	7	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mud/1605.01	1	L,B	L,B	L,B							L,B	L/C,B	C,B	C,B	C,B	C,B	C,B	
Mud/1605.02	1	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mud/1616.01	4	L,1	L,	L,							L,	L,	L,	L,	L,	L,	L,	
Mud/1616.02	1	L,B	L,	L,B							L,B	L,B	L,B	L,B	L,B	L,B	L,B	
Mud/1616.03	2	L,B	L,B	L,B							L,B	L,B	L,B	L,B	L,B	L,B	L,B	
Mud/1616.04	1	L,1	U/U	U/U							C,	C,	C,	C,	C,	C,	C,	
Mud/1616.05	1	L,B	L/C,B	L/C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mud/1616.06	1	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,	C,B	C,B	
Mud/1616.07	3	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mud/1640.01	2	L,1	L,	L,							L,	L,	L,	L,	L,	L,	L,	
Mud/1640.02	1	L,B	L,	L,B							L,B	L/C,B	C,B	C,B	C,B	C,B	C,B	
Mud/1640.03	1	L,B	L,B	L,B							L,B	L/C,B	C,B	C,B	C,B	C,B	C,B	
Mud/1640.04	1	L,B	L,B	L,B							L,B	L,B	C,B	C,B	C,B	C,B	C,B	
Mud/1640.05	1	L,B	L,	L,B							L,B	L,B	L,B	L,B	L,B	L,B	L,B	
Mud/1640.06	2	L,B	L,B	L,B							L,B	L,B	L,B	L,B	L,B	L,B	L,B	
Mud/1640.07	1	C,1	C,	C,							C,	C,	C,B	C,	C,	C,	C,B	
Mud/1640.08	1	C,1	C,	C,B							C,	C,	C,B	C,	C,	C,	C,	

Mouse ID	No. Individuals	d13Mit115	d13Mit17	DNR059	DNR123	d13Mit15	d13Mit133	DNR48	DNR76	DNR101	TetNR5	TetNR10	DNR037	DNR24	DNR003	d13Mit136	DNR16	d13Mit198
Mud/1640.09	4	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mud/1645.01	3	L,B	L,B	L,B							L,B	L,B	L,B	L,B	L,B	L,B	L,B	
Mud/1645.02	1	C,1	C,	C,							C,	C,	C,	C,	C,	C,	C,	
Mud/1645.03	1	C,B	L/C,B	L/C,B							L,B	L,B	L,B	L,B	L,B	L,B	L,B	
Mud/1648.01	3	L,B	L,B	L,B							L,B	L,B	L,B	L,B	L,B	L,B	L,B	
Mud/1648.02	1	C,1	[illegible]	[illegible]							L,	L,	L,	L,	L,	L,	L,	
Mud/1648.03	1	C,1	C,	C,							C,	C,	C,	C,	C,	C,	C,	
Mud/1651.01	1	L,B	L,	L,B							L,B	L,B	L,B	L,B	L,B	L,B	L,B	
Mud/1651.02	1	L,B	L,	L,B							L,B	L,B	L,B	L,B	L,	L,B	L,B	
Mud/1651.03	2	L,B	L,B	L,B							L,B	L,B	L,B	L,B	L,	L,B	L,B	
Mud/1651.04	1	L,B	L,B	L,B							L,B	L,B	L,B	L,B	L,B	L,B	L,B	
Mud/1651.05	1	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,	C,B	C,B	
Mud/1651.06	2	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mud/1675.01	3	L,B	L,B	L,B							L,B	L,B	L,B	L,B	L,B	L,B	L,B	
Mud/1675.02	1	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mud/1683.01	1	1,B	C,		C,	C,	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,B/1
Mud/1683.02	6	L,1	L,	L,							L,	L,	L,	L,	L,	L,	L,	
Mud/1683.03	2	L,1	L,		L,	L,	L,	L,	L,	L,	L,	L,	L,	L,		L,		L,B/1
Mud/1683.04	1	L,B	L,B		L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B	L,B		L,B		L,B
Mud/1683.05	1	C,1	C,		C,	C,	C,	C,	C,	C,	C,	C,	C,	C,		C,		C,B/1
Mud/1683.06	1	C,1	C,	C,							C,	C,	C,	C,	C,	C,	C,	
Mud/1683.07	2	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
Mud/1683.08	1	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mud/1686.01	1	L,1	L,	L,							L,	L,	L,	L,	L,	L,	L,	
Mud/1686.02	1	L,B	L,B	L,B							L,B	L,B	L,B	L,B	L,B	L,B	L,B	
Mud/1686.03	1	L,B	L,B	L,B							L,B	L,B	L,B			L,B	L,B	
Mud/1686.04	2	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mud/1693.01	1	L,1	L,	L,							L/B,	[illegible]	C,B	C,	C,	C,	C,	
Mud/1693.02	5	L,1	L,	L,							L,	L,	L,	L,	L,	L,	L,	
Mud/1693.03	1	L,B	L,B	L,							L,B	L,B	L,B	L,B	L,B	L,B	L,B	
Mud/1693.04	1	L,B	L,B	L,B							L,B	L/C,B	C,B			C,B	C,B	
Mud/1693.05	1	L,B	L,B	L,B							L,B	L/C,B	C,B	C,B	C,B	C,B	C,B	
Mud/1693.06	1	L,B	L,B	L,B							L,B	L,B	C,B	C,B	C,B	C,B	C,B	
Mud/1693.07	5	L,B	L,B	L,B							L,B	L,B	L,B	L,B	L,B	L,B	L,B	
Mud/1693.08	2	C,1	C,	C,							C,	C,	C,	C,	C,	C,	C,	
Mud/1693.09	1	C,B	C,B	C,B							C,B	C,B	C,B	C,B		C,B	C,B	
Mud/1693.10	8	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mudy/106.01	3	B,	B,	B,							B,	B,	B,	B,	B,	B,	B,	
Mudy/106.02	1	B,	B,	B,							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mudy/106.03	1	C,B	B,	B,							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mudy/106.04	1	1,B	B,	B,							B,	B,	B,	B,	B,	B,	B,	
Mudy/106.05	2	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mudy/106.06	1	C,B	C,B	C,B							B,	B,	B,	B,	B,	B,	B,	
Mudy/106.07	2	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mudy/106.08	1		C,	C,							C,	C,	C,	C,	C,	C,	C,	
Mudy/106.09	2	C,	C,	C,							C,	C,	C	C,	C,	C	C,	
Mudy/115.01	1	1,B	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	B,		1/B,		1/B,
Mudy/115.02	1	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	L/C,B		L/C,B		L/C,B
Mudy/115.03	2	B,	B,		B,		B,	B,	B,	B,	B,	B,	B,	B,		B,		B,
Mudy/115.04	1	1,B	B,		B,		B,	B,	B,	B,	B,	B,	B,	B,		B,		1,B
Mudy/115.05	1	C,B	C,B		C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	B,		B,		B,
Mudy/115.06	2	C,1	C,		C,		C,	C,	C,	C,	C,	C,	C,	C,		C,		C,1
Mudy/115.07	1	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
Mudy/115.08	3	C,B	C,B		C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
Mudy/123.01	1	C,B	C,B	C,B							C,B	C,B	C,B	C,B	C,B	C,B	C,B	
Mudy/123.02	1	B,	B,	B,							B,	B,	B,	B,	B,	B,	B,	
Mudy/136.01	1	B,	B,		B,	B,	B,	B,	B,	B,	B,	B,	B,	C,B		C,B		C,B
Mudy/136.02	1	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		1/B,
Mudy/136.03	2	C,B	C,B		C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B	C,B		C,B		C,B
Mudy/137.01	1	B,			B,		L/C,B	B,	B,	B,			B,	L/C,B		1/B,		1/B,
Mudy/137.02	1	C,1			C,		[illegible]	[illegible]	L/C/1,	L/C/1,			C,	[illegible]		[illegible]		B/1,C/1
Mudy/137.03	4	C,B			C,B		L/C,B	L/C,B	L/C/1,E	L/C/1,E			C,B	L/C,B		B/1,C/1		B/1,C/1
Mudy/138.01	1	1,B			B,		L/C,B	B,	B,	L/C/1,E			B,	B,		1/B,		1/B,
Mudy/138.02	1	1,B			B,		L/C,B	B,	B,	B,			B,	L/C,B		1/B,		1/B,
Mudy/138.03	1	C,B			C,B		L/C,B	L/C,B	L/C/1,B	L/C/1,B			C,B	L/C,B		B/1,C/1		B/1,C/1
Mudy/139.01	1	C,B			C,B		L/C,B	L/C,B	B,L/C/1	L/C/1,B			C,B	L/C,B		B/1,C/1		B/1,C/1
Mudy/139.02	1	B,			B,		L/C,B	B,	B,	B,			B,	L/C,B		1/B,		1/B,
Mudy/141.01	1	1,B			B,		L/C,B	B,	B,	B,			B,	L/C,B		1/B,		1/B,
Mudy/141.02	2	B,			B,		L/C,B	B,	B,	B,			B,	L/C,B		1/B,		1/B,
Mudy/141.03	1	C,B			C,B		L/C,B	L/C,B	L/C/1,B	L/C/1,B			B,	L/C,B		1/B,		1/B,

Appendix 3.2

Microsatellites for which primers were tested for polymorphism characteristics in the strains used for the lethal recessive screen. Seven STS markers from the MIT are included at the end of the list to highlight their position relative to the other markers.

Microsatellite ID	Forward Primer	Reverse Primer	Length of Repeat(s)	Product Size (Ensembl v.41)	Start Position (Ensembl v41)	Repeat Type	Source (BAC or other
DNR001	TTTGCTTGATCCCACTCCTC	GCAAAGGTCCCTCTCAGAAAC	20	188	33455042	(TG)n	41L:
DNR002	TTCTGCACCCATATTTTCATTG	ATCAACCTGGGCAAAGACTG	54	176	33537850	(TG)n	41L:
DNR003	AAGTCAAACATGGCTGCTAGTG	GGAGAAAACAGACTTCATTCAGG	53	214	33497327	(CA)n	41L:
DNR004	GTGCACAAGGAAACACCAAC	GCAGTGATGACATCGTCCAG	92	260	33549341	(TC)n	41L:
DNR005	CAAACCATTTATATACAAACATACACC	AAAATTATGGAGCCTCATTTCATC	26	90	33594418	(CA)n	41L:
DNR006	ATGCAATAATGCCAGCAATG	CCTTTGAAAGTAAACCCTGGAC	69	204	33607871	(CA)n	41L:
DNR007	CAGGGTCTGAAGGAAGATGC	TGCTGGGATATATAAGCTGTGG	84	229	33837654	(TC)n	262.
DNR008	TTCAGGAAGCTCCATCCATC	AGGAGTCCACCACAGTCTGC	35	192	30038831	(CA)n	232I
DNR009	GGTGGCTACCTAGAACAGAAGG	CCCCATTCAACACACGTACC	33	202	30027728	(CA)n	232I
DNR010	TCTTGTTGGCATTTTCTTTGC	TGACACCCCCTTCTGATCTC	46	225	30180014	(TG)n	232I
DNR011	CAGGAAACAGCTAAACATCCTC	ACAACTCGATGCCCAATCTC	53	232	30182226	(TG)n	232I
DNR014	GGCATGCAAGTTTCTCACTG	TCCATGGTAGACCCAGCTATG	59	199	30047515	(GA)n	232I
DNR016	GGAGAATTGTGAGGGAATGC	CAGCCATCACATACCACACC	44	172	33910960	(TG)n	262.
DNR017	AAGCAAAGAAACATGGTGTGAG	CACCTCTTCAATGGATTTTTCC	37	219	33817214	(CA)n	262.
DNR021	CTTGTCAGTGCAAAGGTCTCC	GCCACATATACCCAACACTCC	60	194	32435566	(TG)n	43L:
DNR024	AGCCCACTTCAACTCCTCAG	CCCTCAGGGGTACCTCAATC	91	188	32566444	(TG)n	43L:
DNR026	TCTGTACTCTTCTCCCTGCTTG	CACACCACCCTCAAGTATGTG	59	196	32597280	(TC)n	43L:
DNR029	GAGACGGGAGATGACTGGAC	TGGCCGTCCTTTATTAATGC	36	191	31580827	(TG)n	322.
DNR030	TCACAAAGGTTGGCAACAAG	GAAAGCAGGGCTCTCCAAC	0	221	31592377	(CA)n	322.
DNR031	CCATGGAATGCTTGTCACTG	CCCAATTCTGACAGCATGTTC	162	248	31432079	(TG)n	322.
DNR032	CACAAAGACCAGGCAGGATG	GCCAGGGTTTCAGAGTGTG	26	199	31434703	(CA)n	322.
DNR035	AGCCATCAAATACCCTCCAG	TCGAGCAAGGGCAAGATAAG	80	214	31536877	(TG)n	322.
DNR036	GGGCTTTTCATGGTCAGTTC	ACAGAAGGCCCATGTCTCTG	47	204	31569848	(TG)n	322.
DNR037	AACTGCTGGGGAAAGGTAGG	CACGGGGACCATTACTTCC	92	219	31572310	(TC)n	322.
DNR038	TCAACAACTTCAAACCAAGAGG	TCCAGTAGGTGAAAGAGGAGATAG	156	265	31507106	(CA)n	322.
DNR045	CACCATCAAGAACACTGCAAC	GCAGAAGCACCCCATATACC	90	252	26474106	(TC)n	169
DNR046	CCTTTTGAATGGCCAAGAAG	AGCTGAGCTACATCCCATCC	49	221	26474579	(TG)n	169
DNR048	CATGCAGTCAAAACATGCTG	CCCTACACACAAGGTTCAGG	75	203	27243233	(TC)n	189.
DNR050	CCAGACCAAGCAGGGTAGTC	ACCCCAAATTGAGTGCTGAC	97	247	27169530	(TC)n	189.
DNR051	CCAGTCCAAGCAGGGTAGTC	ACCCCAAATTGAGTGCTGAC	69	221	27134518	(TC)n	189.
DNR052	CACTTTCCCTGAGCCTGAAG	ACACACTTGGGAAGCAAAGG	40	217	25333315	(TG)n	19G
DNR053	TTCCCATGAGATGGACACAG	CCTTGAACTTGGTCATGCTG	43	211	25319783	(CA)n	19G
DNR054	CCAGCATTTTCCTTGTACTCC	GATCTTGGCTTGAGCAGTGG	77	196	25256308	(TG)n	19G
DNR055	TCTGGAACTCACCAGCCTTC	TGACAAGGAGGTGCCTACAG	42	192	29356695	(TG)n	202
DNR056	TGTTGCTTCAAATGAGTGTGC	ACAACAACAGGCATGAATGG	30	188	29362754	(CA)n	202
DNR057	TGGCATTCACTGACATCTGC	GGTCATTGCAGAATGGGTATG	45	205	29282143	(TC)n	202
DNR058	GGTTTCTGGAGAACATTTCAGG	GGGGAAGGATTTGTGTTAGTTTAC	72	202	29373595	(TG)n	202
DNR059	GTTCTCACGGCCATTTCAG	CCATGATGTGAGCAAGAAACC	68	199	21786093	(TC)n	38E
DNR060	TAGCTGCCTCTGGGAATCTG	CTGGCACTTGTGTCATCACC	55	220	21784737	(TG)n	38E
DNR061	TTGGAGGTGAACCAGAGTCC	GGACAGCCTTAGACAGGCTAC	58	234	21878734	(TG)n	38E

DNR063	TCTTTCTGCACTGGTGCTTG	GGCCCAACCATCTATCTGTC	51	153	21750467	(TG)n	38E20
DNR066	AACTTGTAGCAAAAGGCTAATGG	AGTGGCCACAGCTCCTAATC	45	210	24421260	(CA)n	80C18
DNR068	CAAGTCCTACACCCCAAACC	CTCCCCTCTCTCGAGTTTCC	40	197	24515259	(CA)n	80C18
DNR070	GAATCCGTTCCCACTCAGTC	TTGCTGAACCTTCATTGCTG	50	194	24502489	(TG)n	80C18
DNR071	GGCTTCTTTCATGCTCCAAC	TTTGTGGGCAGTTTTAGGTG	112	192	24503750	(CA)n	80C18
DNR073	TAAAAGCCCACCACCATCAC	CAGAAGAGCCATTTTGTTGG	47	198	28358882	(CA)n	128M23
DNR074	CAGGGCAGAGAAAACACAAAG	ACCAGTGGGTCCAGACAGAG	26	242	28408701	(TG)n	128M23
DNR075	TATCCTCTGCATTGGCCTCT	ACCTCCCTCTCTCCATCCTC	27	199	28297955	(GA)n	128M23
DNR076	CATTAAGCTCCACCCATTCC	ACAGGAGCAGCTTGTCATTG	104	184	28294295	(GA)n	128M23
DNR077	CCCTGAACACAAACACCAAAC	TTCCCTTTTGGGCTTGAAC	44	244	28423066	(TG)n	128M23
DNR078	AGCACAACCTGGAGAAGGAG	TTATGTTTGGCCAGCTGTTG	43	208	28425336	(TC)n	128M23
DNR079	AACCCAACCCAGCAAAGAG	TGTGGATGTCCCTGGTACTG	56	227	28426337	(TC)n	128M23
DNR080	AGGTTTGACCAAGGCAACTG	TCCATTAAATGCAAGCACTCC	41	201	28431686	(CA)n	128M23
DNR081	CAGGGAAAAGTGGAAAGTGG	TTCATCCCATGACACTCTGC	47	170	28452930	(TG)n	128M23
DNR082	TTGCAAAGGCAATGTAGTGTG	CAGTGGATACCCCACCTTTC	48	144	28472232	(CA)n	128M23
DNR083	TCAAATGTGTGTGTCGCTCTC	CCCTGAGAAATTGGGATGAC	55	204	29612037	(GA)n	47A2
DNR085	GCAAGGGCCTTTCAGATTTA	AGCACCTTCCATTGCTGTCT	42	198	29629330	(TG)n	47A2
DNR087	GCTAGTCCCCTGCCCCTAT	CGGTGCATGGTGTACAAGTC	47	180	29699626	(TG)n	47A2
DNR089	TCTGGAGAGCGACAACAATG	CCTCTGAAACCTTTTCCACTG	51	210	29724871	(TG)n	47A2
DNR091	TGAGTTTGGCATTAGCAAGAG	GTGTGCATTGGAACCATTTG	56	224	29735887	(CA)n	47A2
DNR092	AGGAGTTTGCAAGGGACTTTC	GAGCAATTGATGCTTGAAACC	29	190	29746135	(TG)n	47A2
DNR097	AGGCCAGCCCAGAAAGTC	ATCCACCTCACTGGAACAGG	42	263	29781729	(CA)n	47A2
DNR099	GGAGCCTCTGCTCTAACTGC	TGGGAAAATGTCAGGGTCAC	176	297	29784971	(TG)n	47A2
DNR101	GCAATTAGCAGGAAGGTGATG	CAAAACACCCGAAGGAACAG	70	199	28939517	(CA)n	262c06
DNR102	ACAGCGGGCTGATGATTTTC	GACGAGGTGAGGCTTCTTTC	66	229	29094193	(CA)n	262c06
DNR103	CATGAAAGCATGTGCCAGTG	TTCAATTCCGTATCCCAAGG	37	198	28909972	(CA)n	262c06
DNR104	CTGCAGCATGTTCTCTCTGG	CTCCCCGTCTAAGCACCTC	51	206	29062484	(CA)n	262c06
DNR105	GTCCAATTTGTCTGGACTGTTAC	TTCCATCCTGGTTGTCTTCTG	94	261	29053218	(TG)n	262c06
DNR106	GCAGGAAATTTTCTTACCTTCC	GGCTCTAGACCCTGCTTGC	43	203	29048380	(CA)n	262c06
DNR107	TCACAACCTAGAAGAGATGCAG	TGACAGAGACACCCAAACCTC	25	229	29029004	(CA)n	262c06
DNR108	TCCCGCAAATCATTAAAAGC	TAATTCCCAGGATGGGTCAC	57	213	29022584	(CA)n	262c06
DNR109	CATGGCTCTTACCTCCCTTG	AGACACTGGACACCCACCTC	44	237	29019668	(TG)n	262c06
DNR110	TCCAGAGTGAGACCCAAAGTG	ACAGCCAGGGAGAGTCACAC	84	266	29015610	(GA)n	262c06
DNR111	ATTGGCACTTGCCTGGTAAC	TGACTGTCATGCCTACACCAG	24	208	28968695	(TG)n	262c06
DNR112	CTCCACATCCTTCCTTCAGC	TCTTCTGCCTTGCTGACTCC	43	201	28950732	(TG)n	262c06
DNR113	GCAAGTTCCAAATGTGGTTG	GGTGGGATTGCAAGTCTGAG	24	199	29871682	(CA)n	235O1
DNR114	GGTCTCTCCGGTGGAAATG	AAGGCCATGCGTAAGTGAAG	46	201	29896179	(CA)n	235O1
DNR115	TCTCACCTGGCTATCAACACTG	TGGTGGTGGTTAGAGCCTTC	57	211	29924084	(CA)n	235O1
DNR116	TCTTTACCCCACATGGAAGC	GCTGTGGGGAAGGTGTCTG	47	214	29935651	(CA)n	235O1
DNR117	TAAAGGTGCAGGTGGCAGTC	GAGGGTGGCTTGAGATTCAC	46	198	29968875	(TG)n	235O1
DNR118	GCTAGGGTGTCTTCCAGCAC	AAAACCAAGGCAACTCAACTC	71	191	29974757	(CA)n	235O1
DNR119	GGACAATGCTTACAGCCAGTC	AGGGAGGATGGTAGGAGGTG	45	211	29988457	(TG)n	235O1
DNR120	ACTTCCCAGGCGCTGATAC	AGTCAAAGCACCCAAACCTG	61	227	29988698	(CA)n	235O1
DNR121	ACAAAGGTGCTTCCCCACTC	CAACTGCCAGTGATGGATTG	31	204	23874568	(TG)n	480B19
DNR122	GACTGCAGGGAGCTGTGTTC	TGGGAGCTCTCTGTGACACC	26	192	23851574	(TC)n	480B19
DNR123	GAGGAGTGTATGGCCTGATTTC	CCCAGGTAGATGGAATGAGC	34	193	23849188	(GA)n	480B19
DNR124	AACGCTTTCCTCCCTGAATAC	CCCTAAGGCAGCTTGCTAAC	36	195	23810696	(TC)n	480B19
DNR126	TGGCATGTGTCTCCATTCAG	ATTGCCAGTGCACCTATACC	38	201	23807798	(TG)n	480B19
DNR127	CCAGGCCTATGTGTTTGACC	AATTCTCAAGCAACCACATCC	45	196	23801311	(CA)n	480B19
DNR128	ATGATGGCAATGATGACAGC	CTTTGTGGCTCTGTGTCTCG	99	215	23787001	(CA)n	480B19

DNR129	CAGAGCTGTCCAATGAGCTG	TCTGGCTCCAAACATTTTCAC	51	209	23780781	(CA)n	480E
DNR130	ATCACAAAGGCCCAAGGAC	GGTCCTCCAGCGCATAATC	72	180	23741556	(TG)n	480E
DNR131	CATCTTGAAGCCATTTTCAGG	GGCACTTGAAGGGAAAGAGG	36	198	23713598	(TG)n	480E
DNR132	TTGGCTTGGAGAATTCTTGG	GGAAGTGAGCATGTGCCATC	88	264	23709822	(TG)n	480E
DNR138	AGACTGACCCTGCCTCAATG	CTCATACCTGGCGTGCTTG	88	201	30877721	(CA)n	412F
DNR141	ATGCCTAAGCATGCCCTTAC	CTGCTCTGGCTTTAATGTTCTG	49	196	30921261	(GA)n	412F
DNR152	GCCCTAAACTGTGTTCTATCTGC	CACACAGAGAGGACCTTTGATAAG	82	233	31154531	(GA)n	412F
DNR157	TCTGGGTGGGGTAGACTTTG	TCATTTTGGTGGATCCTTGG	76	197	31201624	(GA)n	412F
DNR158	TCCTGAGCAGGGAGAGATTG	AGCAGAAGATGTCCCATTGC	129	210	31193950	(TC)n	412F
DNR173	AGCAGGCAAGGTAAGAATGG	GCCATTGCAATCAGAAAATACC	120	249	31329775	(CA)n	235j
DNR178	GTGGACAGTCACAGCTCCTG	AAGGGGAGCTGAGAGGAAAG	28	253	24647201	(CA)n	461C
DNR179	TCCCTCTATCCTGGTCAGTTG	CCTATCCATTCCCTGACTGC	129	215	24672611	(TG)n	461C
DNR180	TTGGGCTGAGACACTGAATG	TTGTGGAGGGAGAAAAGGAC	46	218	24691764	(CA)n	461C
DNR181	GCAAAAGAACATGGGATTGC	CATGCACTTTAAGGATGAGTGG	50	202	24708116	(CA)n	461C
DNR184	GGGGCATTCATGTTGTCTTC	GAGGGGAGCTTTGATTAGGG	53	223	25191920	(CA)n	211N
DNR185	CAGTCCTCCAGGCCAATAAG	GAACCTGTGTCAAGGGAGAAAG	62	221	25188742	(TC)n	211N
DNR188	TGGGCAAATTAAACCAAAGC	CTTTTCCACCTGACCCTGAG	58	232	25180028	(CA)n	211N
DNR189	AGCTGAAGCACTGGTGTGTG	CCTTCTGACTTGTGGCATCTC	384	493	25173039	(CA)n	211N
DNR190	AACATAACATGCTGTCTGTCTGG	TCCCTTGTAAGCATCCATCC	82	184	25171464	(CA)n	211N
DNR194	AGGCTCAGCAGGGATTTGAG	ATAGCCAGCAAGGATGGATG	87	228	25138493	(CA)n	211N
DNR195	TTTGCTTTGGGTTTTGTATTG	CCTCACATTCCCCGAGAAG	81	250	25127437	(TG)n	211N
DNR201	ACACACACGGCAAGTGCTC	TTCATGGCTCCACTGTTCTG	224	299	25057837	(CA)n	211N
TNR001	CACCTCAGAGCTGCAAACAC	TGCTGGGTTGTAAGAAATCC	242	395	30085191	(CAG)n	232N
TNR002	TCCAAACAACCAAGAAGCTG	ACCATGAATGTGGCTCCATC	62	199	29310462	(CAT)n	202F
TNR003	GAATGAAACCCTGGCGAAG	TACTCCCTGGTACCCAATCC	105	196	31011250	(TGG)n	412F
TetNR005	TCTCCTGCATGCTCTGTGTC	AACAGCCTGGTCTATAATGCAAG	87	306	30031499	(TTTC)n	232N
TetNR007	CAGCTGAGTGGAACTGAGTCAC	GGGAGGTTGAAGACATGTAGG	111	250	30190056	(CATA)n	232N
TetNR010	GCTGATGCTAGGGCTTTCTC	CCCTAGGGCACATTTCACTG	44	212	30916300	(CGAA)n	412F
d13mit115	TGGTGAAGTGTTTGGAAAAGG	TTTAACCCATTGATCTACTTCAAGG	0	144	20503495	(TG)n	MIT
d13mit136	TTTTATCTATTGAGTAGATTCATGGTG	TATGCCTGGAGGAAAACAGG	0	95	33908480	(TG)n	MIT
d13mit17	CACCCCCAAGTTCTCTTGAA	CCCACATACACATGTGCACA	0	168	21300731	(TG)n	MIT
d13mit133	TAGACACTTAATTCTGTGATGAAATGG	AGCAAAAGCCCCAGTTAGTG	4	150	26234267	(TG)n	MIT
d13mit117	TGGTGACTACCTGATGGACG	TTGCCTGCAAAGTACCCTCT	1	118	37570429	(TG)n	MIT
d13mit15	AGGAACAGCAAGCTCTAAGGG	GGCCTCCCAAGATATCATCA	0	143	24407417	(AGAT)n	MIT
d13mit198	TTCATGAGTCCCTAACCAATAGTC	TCTTAAATAGTCCATGCATTGGC	0	133	34980274	(TG)n	MIT

Appendix 3.3

BAC Position in Contig	RPCI-23 BAC ID	Marker	CA	SP	B6	101	3H1	SA	C3H	BALBc
5	38e20	DNR059			U	U		C;L	S;L	C;S
		DNR060			U	U		C;L	S;L	C;S
		DNR061			N	N		N	N	N
		DNR063			U	U		C;L	S;L	C;S
21	480B19	DNR121	U	U	N	N		N	N	N
		DNR122			U	N		N	N	N
		DNR123	U	U	U	L		U	U	1
		DNR124	U	U	U	N		N	N	N
		DNR126			U	N		N	N	N
		DNR127	U	U	N	N		U	N	N
		DNR128	0	U	N	N		N	N	N
		DNR129	U	U	N	N		N	N	N
		DNR130	U	U	N	N		N	N	N
		DNR131	U	U	N	N		N	N	N
		DNR132	U	U	U	N		N	N	N
26	80c18	DNR066			N	N		N	N	N
		DNR069			N	N		N	N	N
		DNR070			N	N		N	N	N
		DNR071			N	N		N	N	N
29	461o24	DNR178			N	N		N	N	N
		DNR179			N	N		N	N	N
		DNR180			N	N		N	N	N
		DNR181			N	N		N	N	N
32	211M15	DNR184			N	0		N	N	N
		DNR185			N	U		N	N	N
		DNR188			N	U		N	N	N
		DNR189			N	N		N	N	N
		DNR190			N	N		N	N	N
		DNR194			N	N		0	0	N
		DNR195			N	N		N	N	N
		DNR201			U			U		
33	19g04	DNR052			N	N		N	N	N
		DNR053			N	U		N	N	N
		DNR054			N	N		N	N	N
41	169m24	DNR045			N	U		N	N	N
		DNR046			N	N		N	N	N
46	189a16	DNR048			U	U		L	U	S
		DNR050			N	N		N	N	N
		DNR051			N	N		N	N	N
54	128M23	DNR073	U	U	N	N	N	N	N	N
		DNR074	U	U	N	N	N	N	N	N
		DNR075	U	U	N	N	N	N	N	N
		DNR076	U	U	U	N	N	N	N	N
		DNR077			N	N	N	N	N	N
		DNR078	U	U	N	N	N	N	N	N
		DNR079	U	U	N	N	N	N	N	N
		DNR080	U	U	N	N	N	N	N	N
		DNR081	U	U	N	N	N	N	N	N
		DNR082	U	U	N	N	N	N	N	N
60	262C06	DNR101	U	U	U	N	N	N	N	N
		DNR102	U	U	N	N	N	N	N	N
		DNR103	U	U	N	N	N	N	N	N
		DNR104	U	U	N	N	N	N	N	N
		DNR105	U	U	N	U	U	N	N	N
		DNR106	U	U	N	N	N	N	N	N
		DNR107	U	U	N	N	N	N	N	N
		DNR108	U	U	N	U	U	N	N	N
		DNR109	U	U	N	U	U	N	N	N
		DNR110	U	U	N	N	N	N	N	N
		DNR111	U	U	1	B	U	C;L	S;L	C;S
		DNR112	U	U	N	N	N	N	N	U
63	202F03	DNR055			N	U		N	N	N
		DNR056			N	U		N	N	N
		DNR057			N	U		N	N	N
		DNR058			N	U		N	N	N
		TNR2	U	U	N	N	N	N	N	N
65	47A02	DNR083	0	0	N	U	U	N	N	N
		DNR085	U	U	N	U	U	N	N	N
		DNR087	U	U	N	N	U	N	U	N
		DNR089	U	U	N	U	N	N	N	U
		DNR091	U	U	N	N	N	N	N	N
		DNR093	U	0	N	N	N	N	N	N
		DNR097	0	U	N	U	U	N	N	N
		DNR099	U	U	N	U	U	N	N	N

BAC Position in Contig	RPCI-23 BAC ID	Marker	CA	SP	B6	101	3H1	SA	C3H	BALBc
66	235O1	DNR113		U	N	N	N	N	N	N
		DNR114		N	N	N	U	N	U	N
		DNR115		U	N	U	U	N	N	N
		DNR116		0	N	U	U	N	N	N
		DNR117		U	N	N	N	N	N	N
		DNR118		U	N	N	N	N	N	N
		DNR119		U	N	U	U	N	N	N
		DNR120		U	N	N	N	N	N	N
67	232m10	DNR008	U	U	N	U	U	N	N	N
		DNR009	N	U	N	U	N	N	N	N
		DNR010	U	U	N	U	N	N	N	N
		DNR011	0	U	N	U	N	N	N	N
		DNR014	U	U	S	U		B	L	C
		TetNR5	0	U	L	U	U	U	U	B
		TetNR7	0	U	N	U	U	N	N	N
		TNR1	U	U	N	N	N	N	N	N
		TNR3	0	U	N	N	N	N	N	N
73	412P21	DNR138	U	U	U	3	1	C;L	S;L	C;S
		DNR141	U	U	N	N	N	N	N	N
		TetNR10	0	U	U	U	U	C;L	S;L	C;S
75	408H17	DNR152	U	U	U	U	U	C;L	S;L	C;S
		DNR157	U	U	U	U	U	C;L	S;L	C;S
		DNR158	U	U	U	U	U	C;L	S;L	C;S
76	235J20	DNR173	U	0	U	U	U	C;L	S;L	C;S
78	322j11	DNR029	U	U	N	U	U	N	N	N
		DNR030	U	U	N	N	N	N	N	N
		DNR031			N	N	N	N	N	N
		DNR032	U	U	U	U	U	C;L	S;L	C;S
		DNR035			N	U		N	N	N
		DNR036	U	U	N	U	U	N	N	N
		DNR037			U	U		U	U	U
		DNR038			N	N		N	N	N
86	43L14	DNR021			1	B		C;L	S;L	C;S
		DNR024			U	U	N	N	N	N
		DNR026			N	N	N	N	N	N
93	41L21	DNR001	U	U	N	N	N		N	N
		DNR002	U	U	1	B			L	C
		DNR003	U	U	1	B			L	C
		DNR004	U	U	1	B			L	C
		DNR005	U	U	N	N	N		N	N
		DNR006	N	U	N	N	N		N	N
95	262j21	DNR007	0	U	1	B		C;L	S;L	C;S
		DNR016			1	B		C;L	S;L	C;S
		DNR017			1	B		C;L	S;L	C;S

Key to data for appendix 3.3

Polymorphism data for each of the tested microsatellites identified in the del36 region.
Green = shared product sizes with 2 or fewer other strains (B)= c57BL/6J. (C)= C3H. (1)= 101/H. (L)= BALB/c.
Red (U) = Unique product size. White (N) = Not informative – shared product size with 3 or more other strains.
White (0) = Product failed to amplify. Grey = Not tested.

Appendix3.4

Del36H Hybridisations: Singletons.

NB: Not every probe set produced singletons. Of the single probe hybridisations, only Gpx5Int1 and Spi3 resulted in singletons.

33 Overgos

100g12	100h3	101h21	102i6	105j10	105k17	105m8	105n12	105p17	106a17	107g2	107g4
109a3	109d1	109j1	109m17	110b20	110j6	111e13	111i17	114i6	114p18	115n14	116a7
116g3	116o19	117m10	119b6	122f20	123g19	123n2	124k6	125b5	126g17	126j9	128d15
128n21	129k17	12c24	12o21	130b18	130g10	130l9	131n5	133e17	133g2	133k7	135d8
135d9	136j21	138f15	139g1	140k14	142g18	142j11	142l23	142m7	142n3	142n5	142o8
142p6	146e10	147b15	147c15	148m15	150a5	150d9	152d4	153m5	159i15	159j15	15n1
160g20	162f24	163e19	163o6	164d11	164e18	166a16	166g18	166i16	167c15	16k11	170l6
171e12	171i4	171m8	172e10	175b1	175b18	175l5	176k14	178f21	17o12	181k16	182i1
182k23	184b2	186o2	187h1	187n9	18b10	18h3	194k8	194p5	195h15	195l9	196a16
197h6	198c15	199d13	199d7	199o23	19d11	19d17	1a17	1d14	1o16	200a14	200d22
200p6	202e11	202e17	202f3	205b16	206e11	207j21	208f8	209o5	209p4	20d17	210a15
211o18	212n16	216a19	216c22	216c23	218g4	218o3	21a9	220a19	220b21	220k8	220m23
222e18	224e14	226f16	22e13	22o13	230d9	230j5	234g21	235o21	237j24	23b21	23f16
242j12	242l22	242m1	244f22	244p8	251m12	253b15	254m3	254n9	259h20	259i19	25k6
25p12	260l12	264n17	266h16	267d22	26a14	26f15	26h3	273o7	274b5	276i19	277f11
279g23	27l7	281e23	283a12	283b8	285a21	288b12	293l24	295i19	297b23	29p5	302l3
303i12	303j12	304h1	305a6	305m20	30a12	310d18	311h17	311i4	314k8	317a19	317p13
321p1	322f18	326h6	329e9	329g10	330c17	333j6	336i12	336k18	33f22	340a6	341a9
341p22	345d8	346p6	348h2	34l10	350h12	350k4	351k21	356o6	360b12	360b7	363i21
366f1	371g9	373g9	376l17	378c21	378f18	37c8	381f15	381i15	381n18	383g23	385i12
385m12	386j6	388a11	391j15	391n15	394d8	395a9	395p8	398g13	398p4	39j20	401c18
402c9	403b15	403d13	404h14	405k19	406n18	408l10	411n9	413j16	421c9	422b20	425k8
425p16	426l8	42b5	432e19	432f11	433c15	433f8	434m5	434o11	435l14	435o20	437i2
438j10	438j22	438k22	441n20	442f16	442m7	443d20	443l6	444l19	446o22	447e23	447g4
447m15	448e4	448o21	449e11	449h14	449n17	450i7	451p22	453i11	454d20	454e19	455b13
456a20	456e19	456n10	457j21	458l10	459l8	459m10	459o9	461e24	462m14	465p6	466b3
466k12	466k6	466l13	466m14	467l17	469d23	469l6	469o13	471o18	472e8	472p1	477k19
478n24	479d22	479n1	47f14	50n14	51d10	52c11	52m8	53b20	53f23	53l15	57f5
57n23	59a19	59b21	61l21	61p4	62h20	64c7	64f21	68l12	6e16	70n11	71b1
71f7	73f14	73n20	74b12	74c11	74m13	75k14	77g6	77j15	78g13	78n9	79n23
7g7	7j19	7j4	80e17	80f16	82f19	82h20	83k21	84m22	85e12	85k19	85k20
86l2	86m15	86o24	87b12	87b6	88f8	8c15	8i10	8m9	91e15	91p12	92e21
92i6	92j22	92k16	92m6	92o18	93j19	94j6	95d3	96e9	97g14	97g8	97h14
97i15	98e3	98g3	98k10	9b15	9l1						

EST set

100i4	100n19	100n2	103i17	104p6	106g11	108e8	10g5	110k10	114n1	117c5	117d7
118f11	120n19	123i13	126c19	127a9	128j18	128l18	130i12	130l20	130m22	133h14	135b15
136o2	138j12	13e12	13o15	141b7	144f23	144g23	144k23	144n2	144o22	145j6	146o23
148e17	149c20	150c22	151h1	152b6	153k6	155p19	164p12	165c7	168f14	168i3	169j18
169k7	170e3	173f7	178e4	179p1	184d3	184l16	186b12	186b4	186e3	186k2	187j17
18c1	192j4	193f8	194b9	194o4	198p4	200o11	201h18	201i4	206d23	206e23	206n5
209n13	210h12	214j20	220n5	223a17	224a19	225i6	22c17	23c8	248b12	250i8	251k1
257f18	261p7	263l11	268e21	268l1	279l10	27d15	27p10	280d24	280j1	280l9	281b22
282g21	283h10	288n20	289n20	28l24	28m5	28p4	296n12	297c6	302n20	303i9	305e16
308b24	308c23	310h13	32j14	339b7	345e12	34m21	350d15	351m2	353g22	362m4	36i18
372e3	374o23	380j6	386b15	386o18	389b1	38h24	391k8	394j8	398d6	402e9	405p10
407a13	409e8	40c18	40g20	416o10	418e8	422h17	425d17	426a11	428e13	42e21	42p8
432l22	434c17	434m10	435m19	438a4	438f9	440d8	442f21	442g21	443b8	443n16	444c5
445h14	451a14	451b2	454g19	456l7	45i19	45o9	461m4	463l4	466h4	466p16	468f22
468j22	46b21	470k14	474c4	475e11	475n24	48i9	53h5	5d22	62a6	63k6	6j1
76a17	7c5	7k16	7p18	80f8	80g5	80g7	8b19	8l10	90j1	98d11	98n22

FS1

182e20	21g6	262p23	273c19	274c24	294h4	324e24	329f5	330c3	442j23	98k14	

FS2

121o8	127k15	127m15	140d13	22d23	258d13	259i8	267b16	271k15	299l12	314i19	322g6
327a9	362m14	379b19	392j16	395d22	430e6	477i21	479p2	70o23			

FS3

112n22	112p20	120i8	131l22	135f8	287m6	293a6	479h23	57k21

FS4

129d6	130j9	134a8	135c16	135f20	135j16	137n6	164d15	169m4	186i8	204i17	224c20
233f9	248f7	292e8	305g17	320i13	321l17	325n7	326i15	329b12	332i12	332p1	333l21
362p10	363o18	366p2	373j13	375k20	377e3	380i11	390d16	395i21	405c5	412g19	412g20
413a20	416p24	417e17	423h10	424m11	426g2	432g16	442p10	443k19	450h12	453f12	453m24
455d7	456c9	457o17	461p6	469a19	476i17	478h8	479p5	51i24	66k9	70k11	70k19
71n16	80l7	87d22	90k6	92k2	93c1	97o10					

FS5

116e20	136p2	194i12	1e14	209o24	212g5	227a17	245m1	250p9	264a15	272i10	278i1
284l8	287i18	300m6	302p20	309e15	315e23	366k14	370o24	375e13	405o13	427h3	443m5
472o17	476p1	477m22	477n20	80h11	98h2						

FS6

104b16	115d13	121c5	128m23	130b3	133h4	133p4	143o13	155m9	176a14	176b14	176j14
190l8	19h20	212a15	225n9	237i3	247p24	248j6	287k23	296m6	370e23	388l2	40l16
42g13	46c20	471n2	475f17	475o24	53e4	55m16					

FS7

12p11	130o10	135c18	138f20	142l17	148g3	154g4	154k16	156o17	160b12	220h15	236g2
266j1	26j6	280k24	287p20	28h10	28j7	28l5	298f22	29i21	302j9	30d8	30g17
314k21	32j22	341a8	343h24	346c14	347c10	347k3	349f19	350a11	356a14	357d18	358g20
366e3	366k7	372h19	375b8	380l6	385e17	392o20	397e13	423b15	427h7	432b5	435a1
472o18	477c1	47p20	51c6	56m13	61d10	6a5	73d9	74p19	76g13	76h13	94a1

FS8

101l1	109j8	120d8	120l9	121j15	127i9	129c8	129f8	132k11	132l12	133j17	138d7
138f16	140e20	142g8	143h4	147a15	147f21	149f6	151e4	151g23	152a24	155p16	158o17
161d1	161g20	162i7	162k7	164o8	166c2	182g16	183a6	183d11	18j3	196k6	197i9
199c11	208l23	209c9	209g11	210k10	211c15	212b16	216n4	220n18	226b1	230c21	233l17
237a8	241c10	260l10	261i9	262j16	275d13	275d14	277a16	280g12	282a18	291c11	291g9
292n3	294g8	296f8	305o6	306h23	307e5	309o15	310g18	310m18	321g8	323o20	324m3
326a23	328o21	32l7	331e20	332d7	332e6	336i23	344g16	344j2	34e2	353a9	355f13
356b8	357l14	359m8	360m12	370i11	373b6	373p18	374a15	374l15	37d20	381p23	381p8
385b7	385h6	389c22	38b1	38g17	38j17	39g24	3c3	402a3	402j14	404n13	417h2
418d19	418f16	421b7	422e24	425k15	429l8	42i20	430n2	436f8	436p10	446m24	44a16
455n24	457n22	463p9	46i3	476n4	480b2	48e8	48k5	49c5	49g7	49k15	49n21
49o15	49o2	63o16	65o16	66c20	6h7	71f19	71m5	78g21	93k7	98c17	98g13
98n14											

FS9

101h16	102k1	103m20	106i6	106m6	116k7	117b10	120m8	120o10	12d8	140a24	144i3
164i3	164l5	178a14	181f20	185e9	196f10	197m12	198a24	202d2	205c4	208a19	210h10
242m4	244k24	251f7	253b7	255h13	256k13	270p4	278k20	290c6	294n15	298c3	302j1
305d11	306d6	307i9	318c24	318d24	318e24	325f13	337h17	339a19	340o6	342p22	343i21
344p18	346k16	349e23	349p1	352o6	353e8	356a6	356c6	361n20	362k19	366m23	367b14
373c9	374g14	376f11	382c1	385b24	388d4	388f4	388j3	391k5	391p2	393g6	399l2
3o7	401f21	401i18	404b22	406n7	406o6	408n3	408n4	408p14	411d17	411o13	411p14
417d8	419e4	420h15	425n1	428k4	430n7	431e20	436m11	440g7	440h7	441c13	447p4
448j6	453g18	453i17	460o3	462c11	462i17	462k20	468o1	469c2	474e17	474h22	474j24
474k20	474k23	474l22	476c2	56g21	75f11	83g22	87o13	97c22	97f20		

FS10

104g3	105e18	132l13	132l14	137o5	149e10	150o19	154h1	167f18	178f3	212d7	212n5
214m9	267m13	270o3	289f15	306g1	311c20	311e19	316a10	317l8	329h18	343j23	346j21
35k21	36g9	394o5	400h23	402d23	411o14	424h22	424n22	440e19	441i3	442d2	452b14
458g23	459g15	468j13	470a15	66m17	70g21	84b3	84i12	85p24			

FS11

109c19	138o19	165n10	202m14	246f16?	254k22?	261g17?	261m11	263m19?	265o1	267p21	270l4?
273m12	274a22?	274a24	279f22	279h22	288a24	288b7?	288g12?	306k14	318n16	342f12	369d2
36j1	375o14	37j12	445j17	449o5?	44j2	454i11	470a8	59m1	77b5	99a10	

FS12

175g12	201k21?	204m15?	208c8	229e22	249e21	276e19	276e20	276h16	277c6	307k11	332p21
374g5	384f18	394l4	407a19	425a7?	426j6	439f14	470c1				

FS13

10n19	129h1	134a1	149b19	166p18	180e4	184m18	192d10	197n11	207b16	226h11	231m4
321f5	335g7	338i7	33p17	343j16	358p9	400p14	412m14	412p21	421o1	421p1	451a10
468m24	4o14	61f2	61j3	62m10							

FS14

10p2	11e9	151o23	152k6	152n6	15b24	193k2	193m3	232l16	287d13	305a17	332f15?
333j18?	335g12?	338p9	394j14?	397b22	414e19	417o14	45h23	469j19	472d13	51a11	51c8
5m15	5o15	62l15	84k13	90e8	93c4	93k5	94b10				

FS15

11c16	121m11	140e1	15j3	170b5	170e9	191b15	205j10	205j13	224n13	250c18	255h14
265n23	268m17	282c8	286h20	2i4	2o11	305f1	305k6	319e20	322i5	329e10	329m10
32c7	339b12	344b10	36h19	370h4	385e19	398c19	3i10	402o2	405p1	407k20?	40g6
443f23	454n1	457g3	459k15	461l21	46i21	476c8	478j6	49p10	61a9	62d15	62m18
66f7	67g3	70b23	8k9	94e12	95k11	97d17	97e10				

FS16

115h3	117d3	338p21	378c24	378d22	378d24	378e24	37b23	38p3	409d23	412i24	415l7
418g9	439i10	50a2	50g2								

FS17a

107n16	108k7	109l5	118f13	118h13	11e10	135a24	135g11	140p5	141h10	147c4	147f5
149k16	151l13	154m14	156j6	157g16	160k13	16j3	171d14	173n11	179m18	183o5	184j8
189e16	189k12	190o2	194a19	198i19	205f11	20a22	210g22	221e16	221h15	222b3	227n16
228b17	229l24	230f9	235j20	235o6	236p21	237g10	239a23	23m9	242p24	246j6	249h22
249h3	251h21	252g12	256k19	256p19	25j15	260d8	262h17	275g21	282o9	287p11	292i19
302d11	305g14	331i20	339k3	344l4	348i13	377b13	37l21	380i8	381d18	387h21	388o1
38l5	401c5	409m10	414j19	419b17	432n16	432n9	432p2	435f10	438a9	43i21	450o2
461p13	52a6	58h13	61k6	67n7	6j22	73b5	73i10	73k3	7h15	86i7	89p15
92g5											

FS17b

119h16	138f6	144m11	144o9	154i9	16d17	185i8	19p19	227p13	241i15	247b18	250h14
258b9	25m17	263d15	263e13	266p4	267n4	267p5	277b9	286i8	303d18	304b10	304d12
321f1	321i5	322g24	326c24	326k20	330d1	332f12	333h4	340h12	354g11	359c1	365h19
371g5	371n20	374m7	395o13	397p19	406d19	406d22	416j12	424p4	426h4	430i14	435e23
442c19	448d21	448l4	451g23	452l13	459c21	459d10	461m24	465i8	465o20	46j24	470n14
471j24	472d15	476m17	477c16	51g20	59c13	60j17	78n15	82j4	9o13		

FS18

125j3	133i2	200c1	230m8	239d22	239f20	239o18	243p18	308o23	326i1	334p2	364k16
369d24	372b7	373o22	377a15	3b5	447h21	471o20					

FS19

241n10	255b9	274p18	302m2	406b5	94n12						

FS20

151a11?	154f7?	163d13?	16a10?	180c14?	18o1?	18p1?	1o8?	413m19?	41p24?	54c23?	54f13?
5d3?											

Gpx5 Int1

267h7											

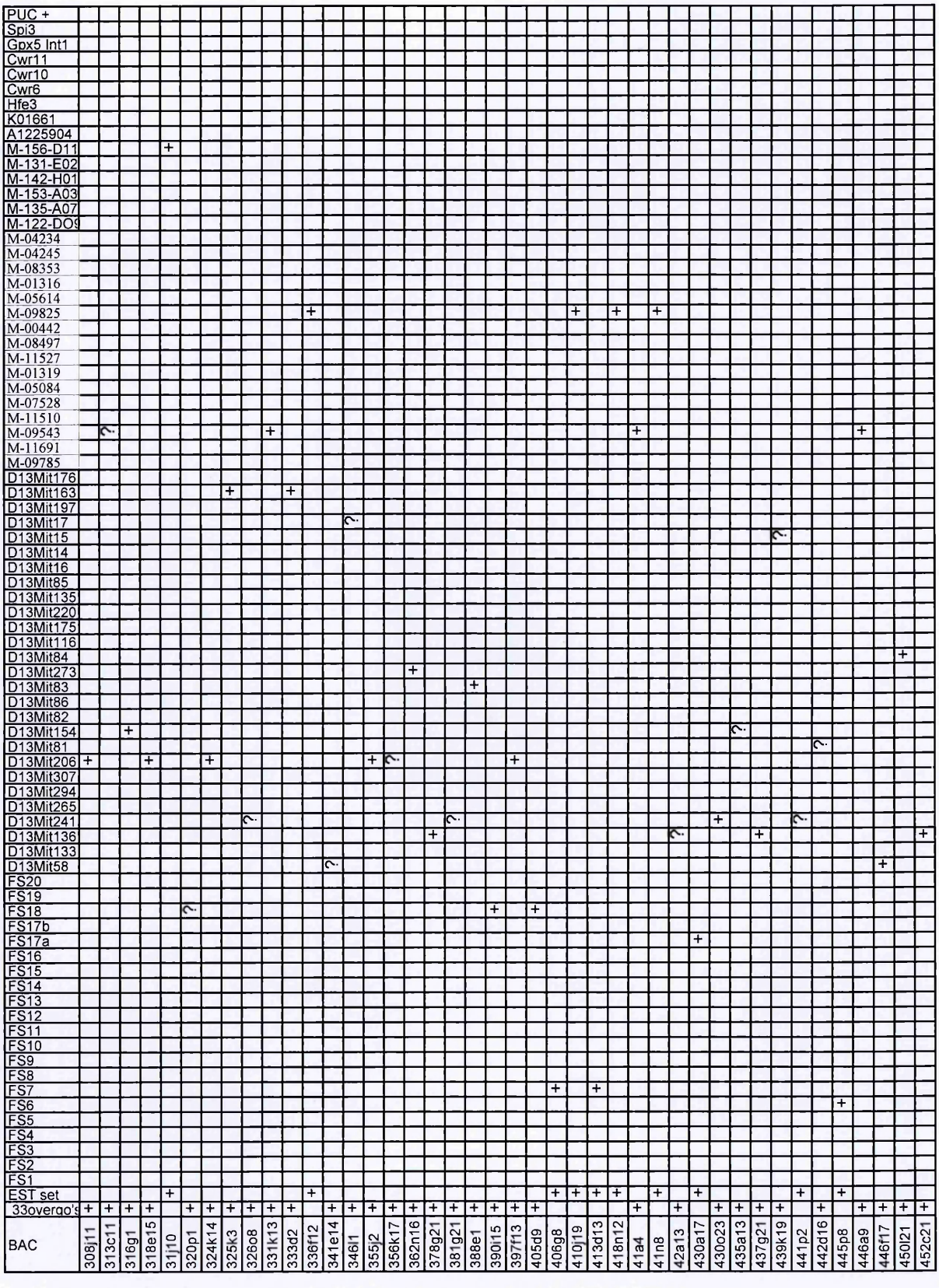
Spi3

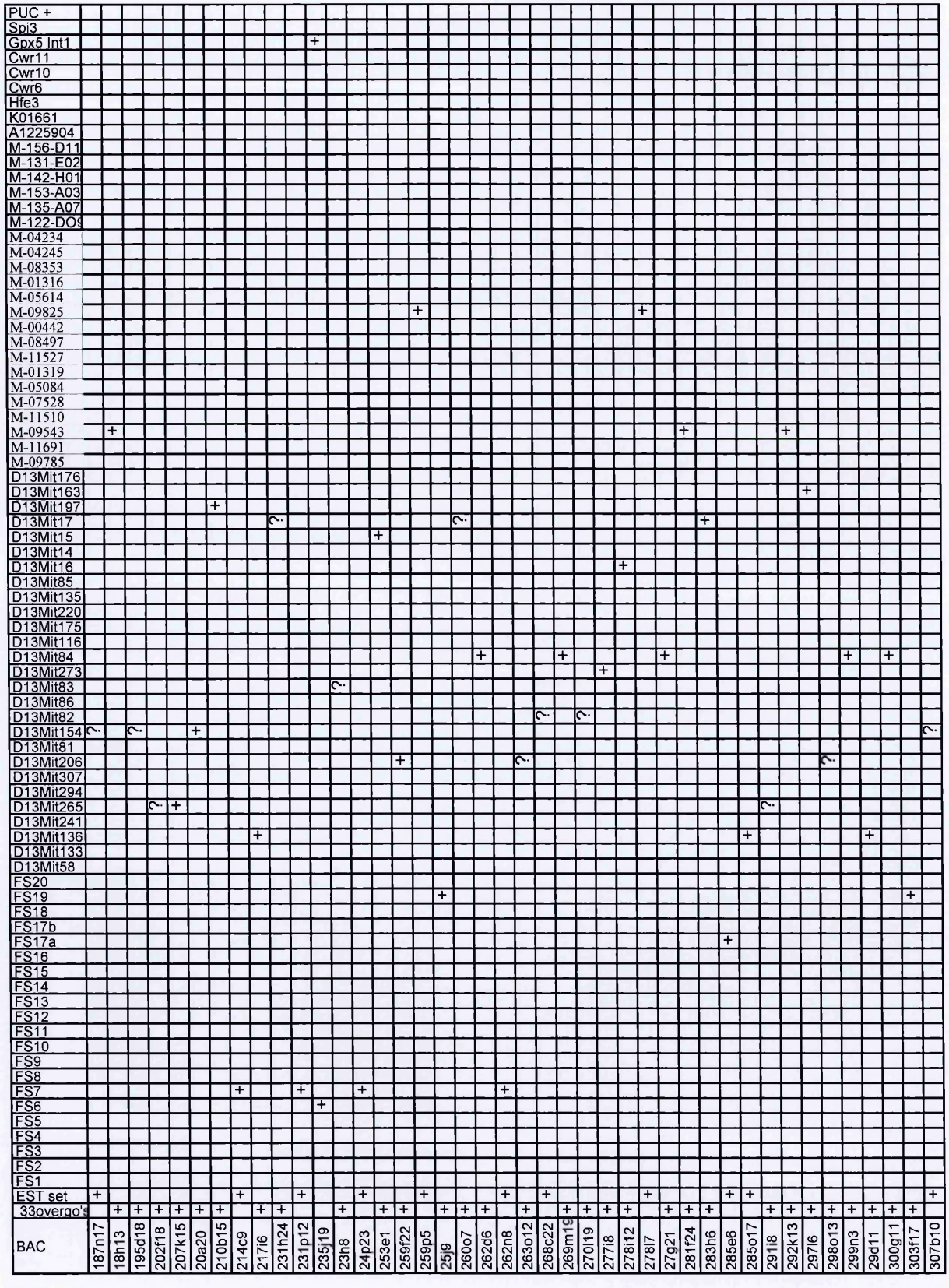
151g4	151o12	164h2	236c1	302f7	305j4	305o4	307o21	330m3	330m4	354h10	369d15
36j15	398a8	403o7	422p14	439e19	450c17	81f1	91p6	94b23			

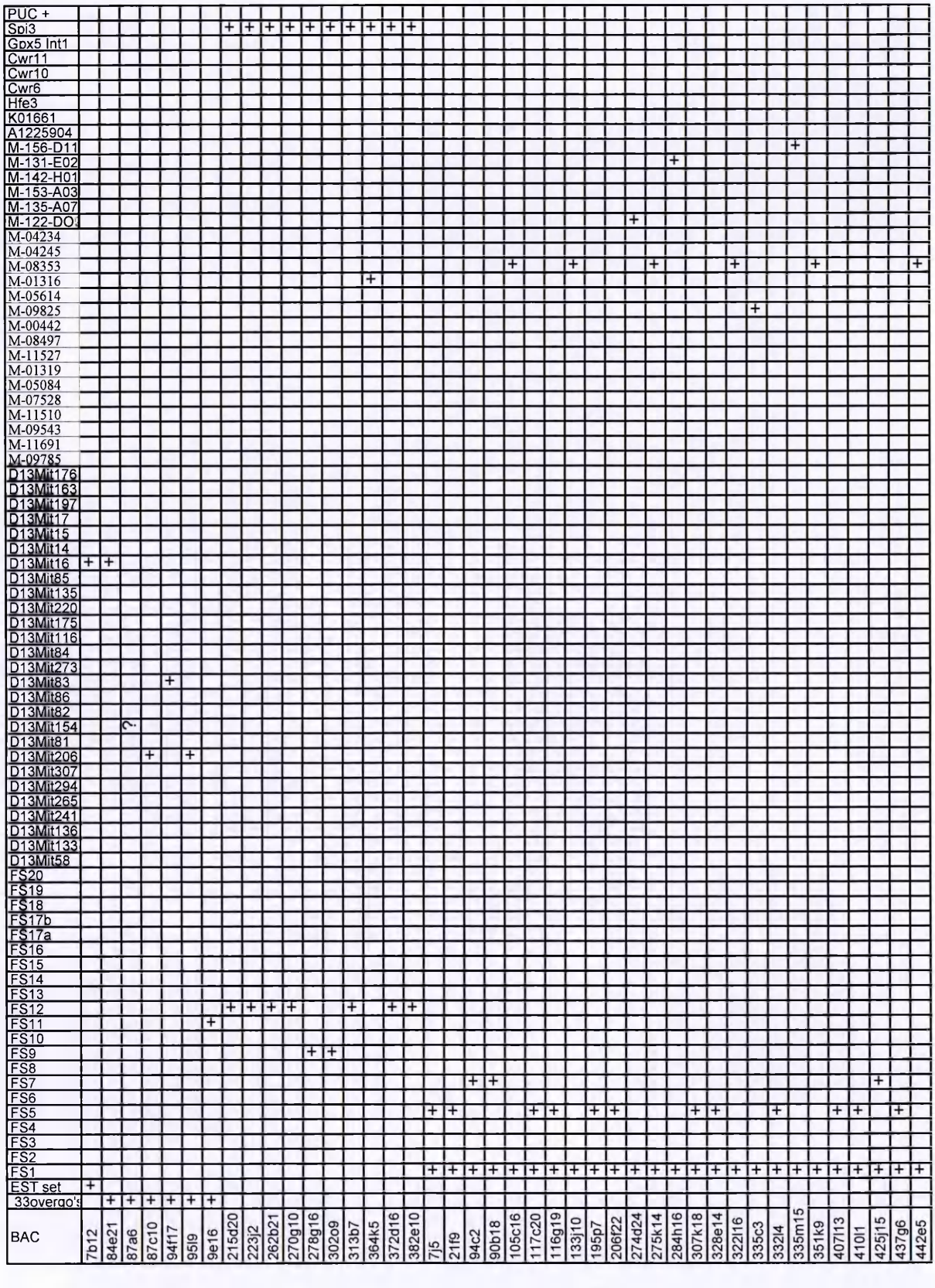
Appendix 3.5

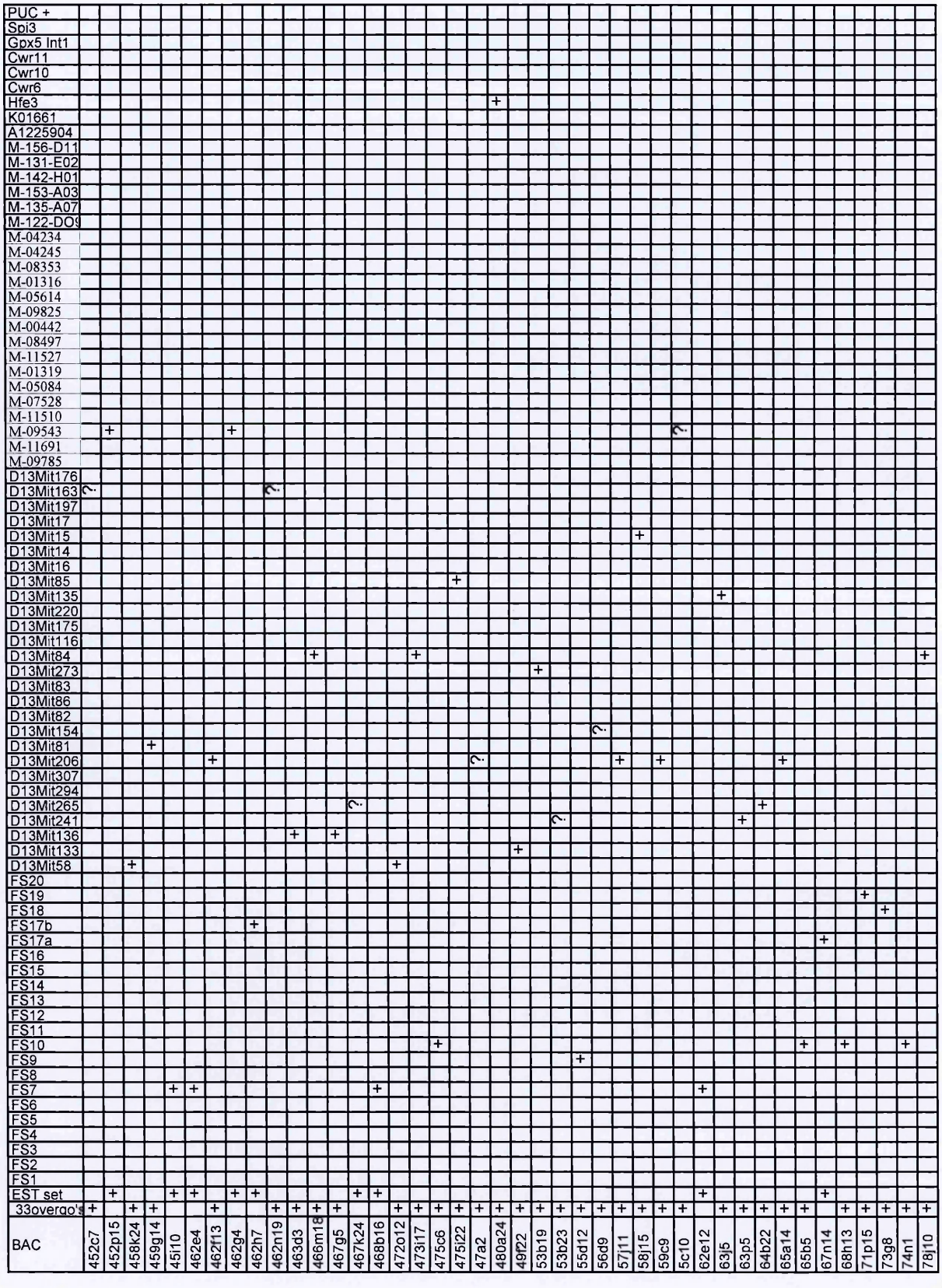
Del36H Hybridisation: Multiple Hits

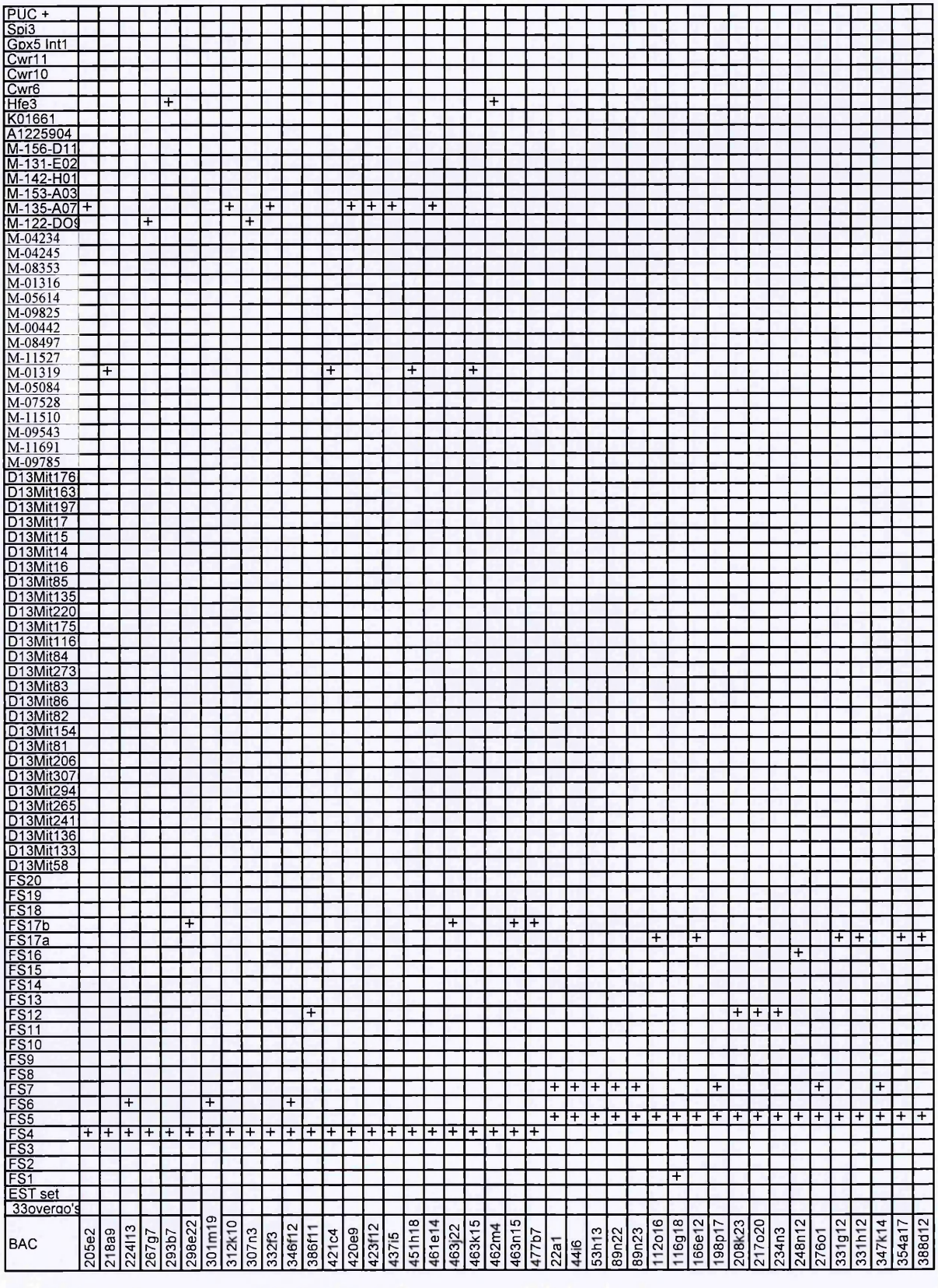
| | Multiple Probe Hybridisations | | |

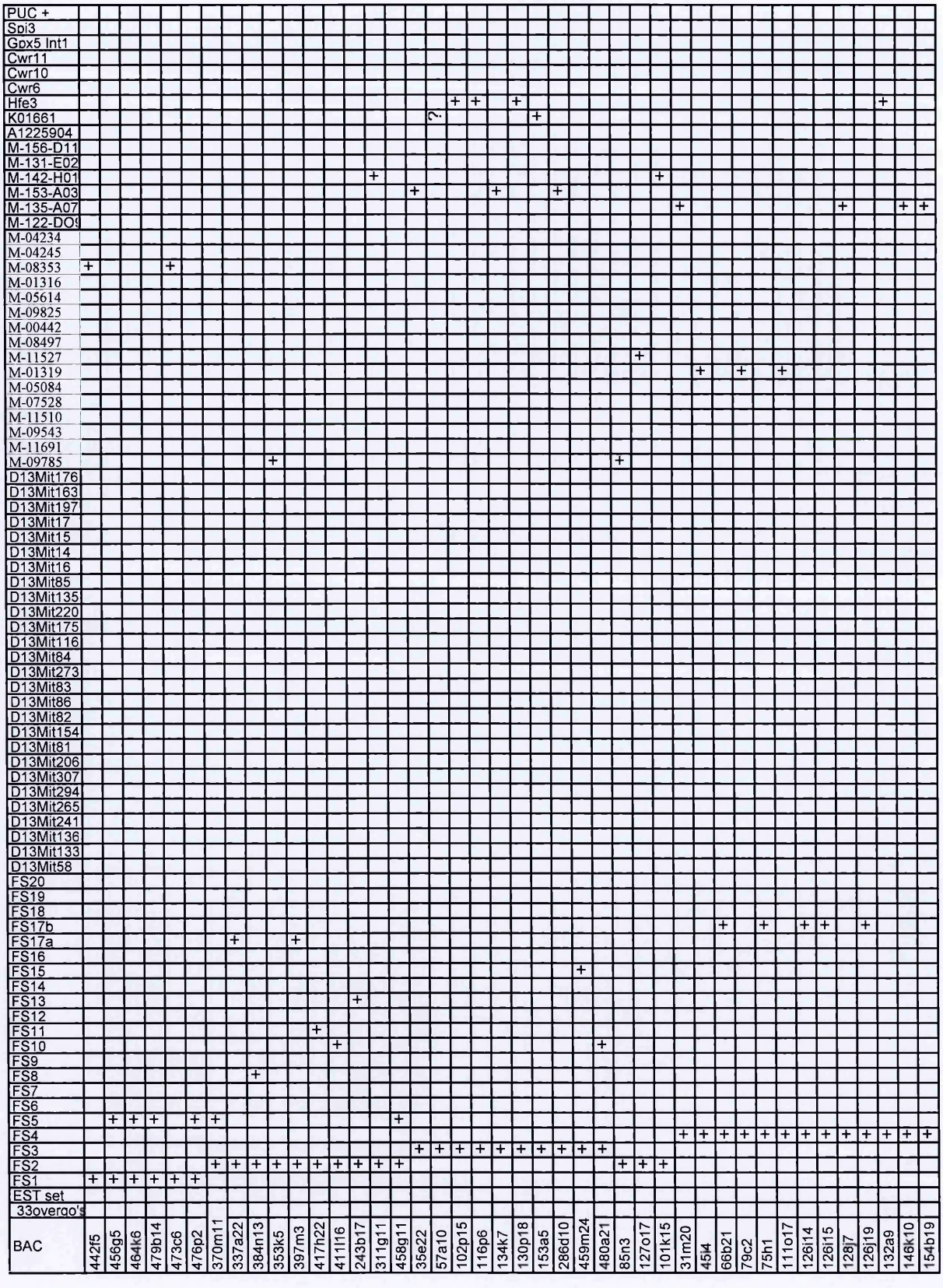

PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-DO9
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
308j11
313c11
316g1
318e15
31j10
320p1
324k14
325k3
326o8
331k13
333d2
336f12
341e14
346l1
355j2
356k17
362n16
378g21
381g21
388e1
390i15
397f13
405d9
406g8
410j19
413d13
418n12
41a4
41n8
42a13
430a17
430o23
435a13
437g21
439k19
441p2
442d16
445p8
446a9
446f17
450l21
452c21

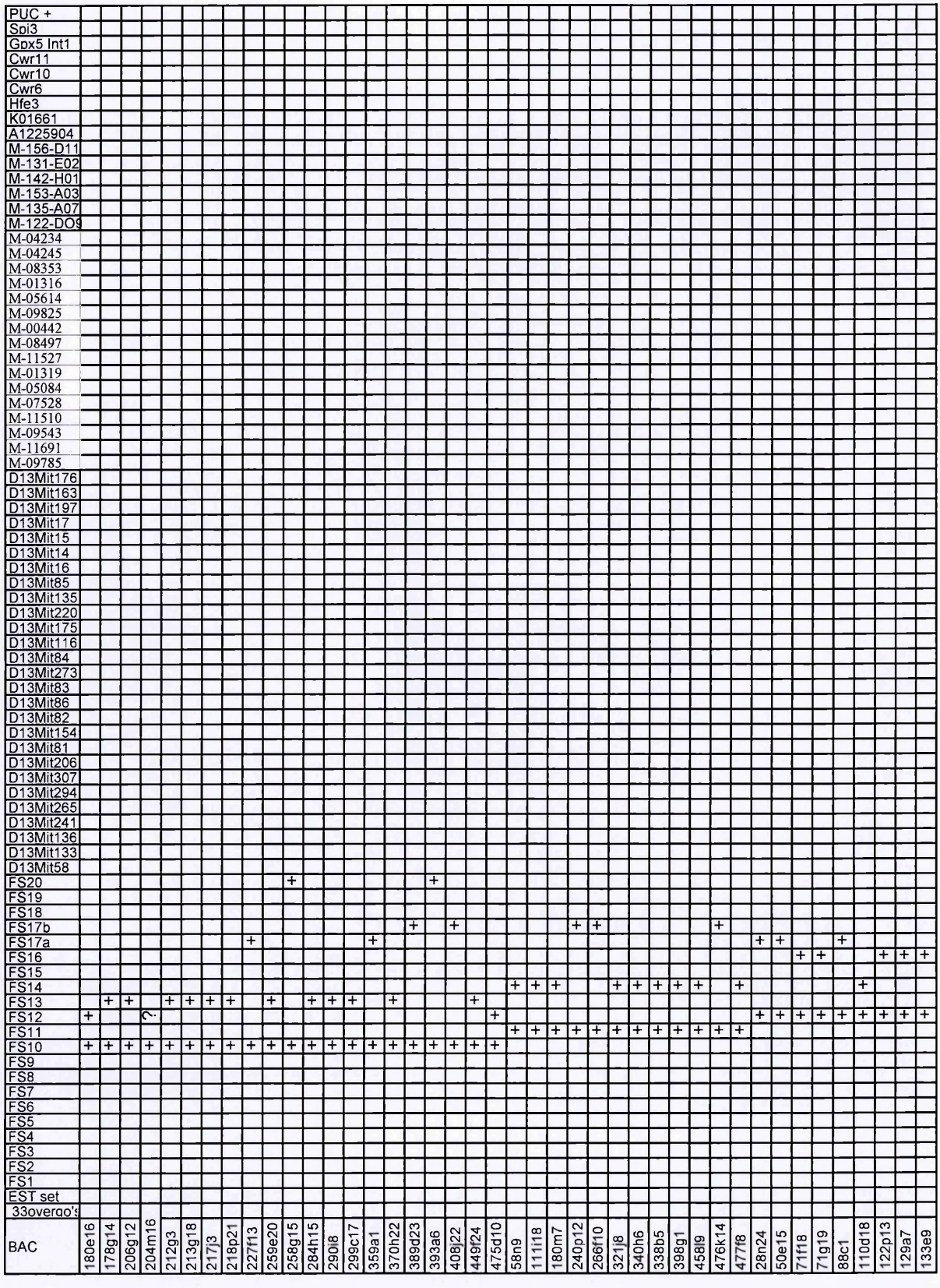

PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-DO9
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
187n17
18h13
195d18
202f18
207k15
20a20
210b15
214c9
217l6
231h24
231p12
235j19
23h8
24p23
253e1
259f22
259p5
25j9
260o7
262d6
262n8
263o12
268c22
269m19
270l19
277l8
278i12
278l7
27g21
281f24
283h6
285e6
285o17
291l8
292k13
297l6
298o13
299n3
29d11
300g11
303f17
307b10

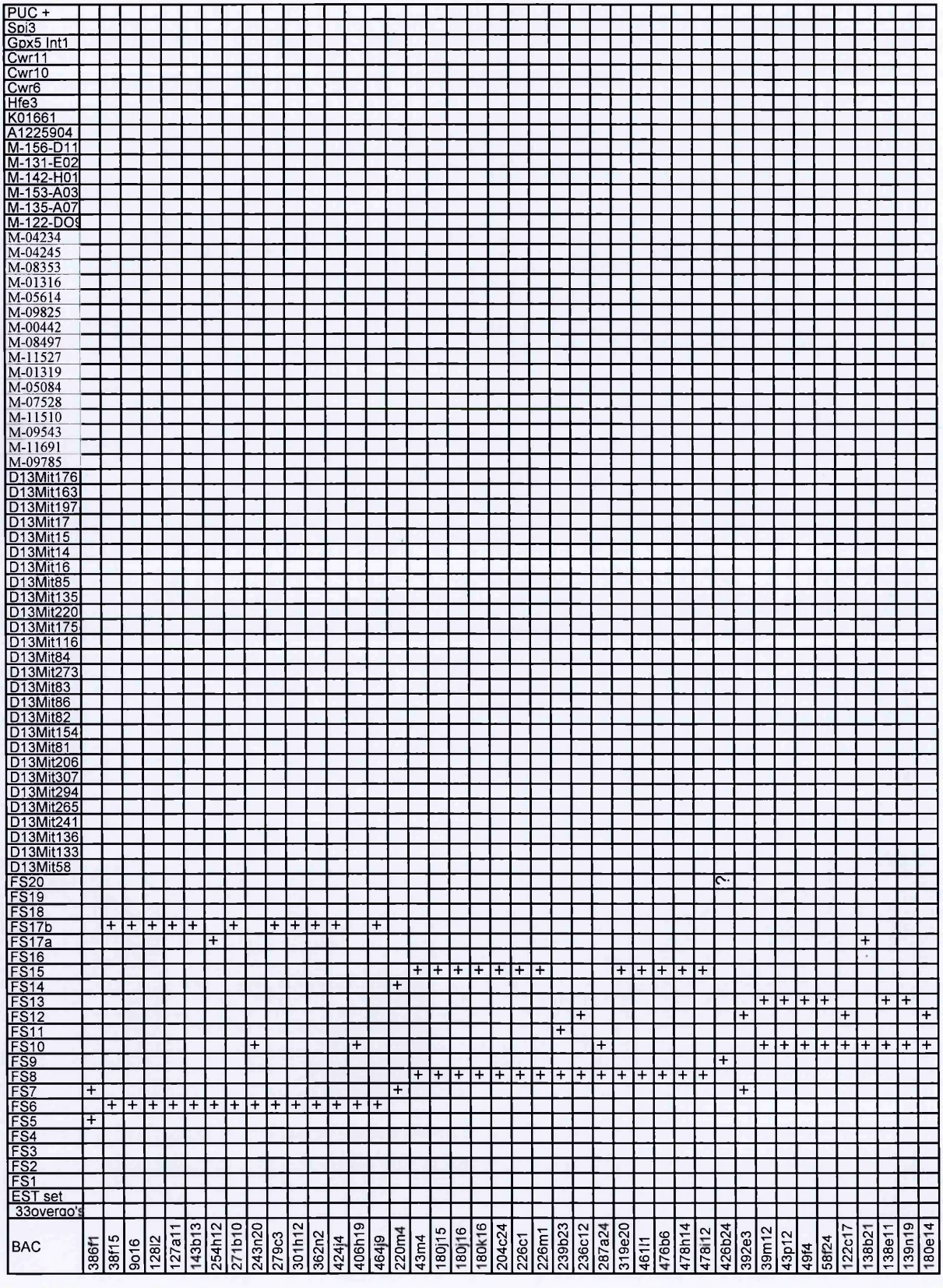

PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-DO1
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
7b12
84e21
87a6
87c10
94f17
95l9
9e16
215d20
223j2
262b21
270g10
278g16
302o9
313b7
364k5
372d16
382e10
7j5
21f9
94c2
90b18
105c16
117c20
116g19
133j10
195p7
206f22
274d24
275k14
284h16
307k18
328e14
322l16
335c3
332l4
335m15
351k9
407l13
410l1
425j15
437g6
442e5

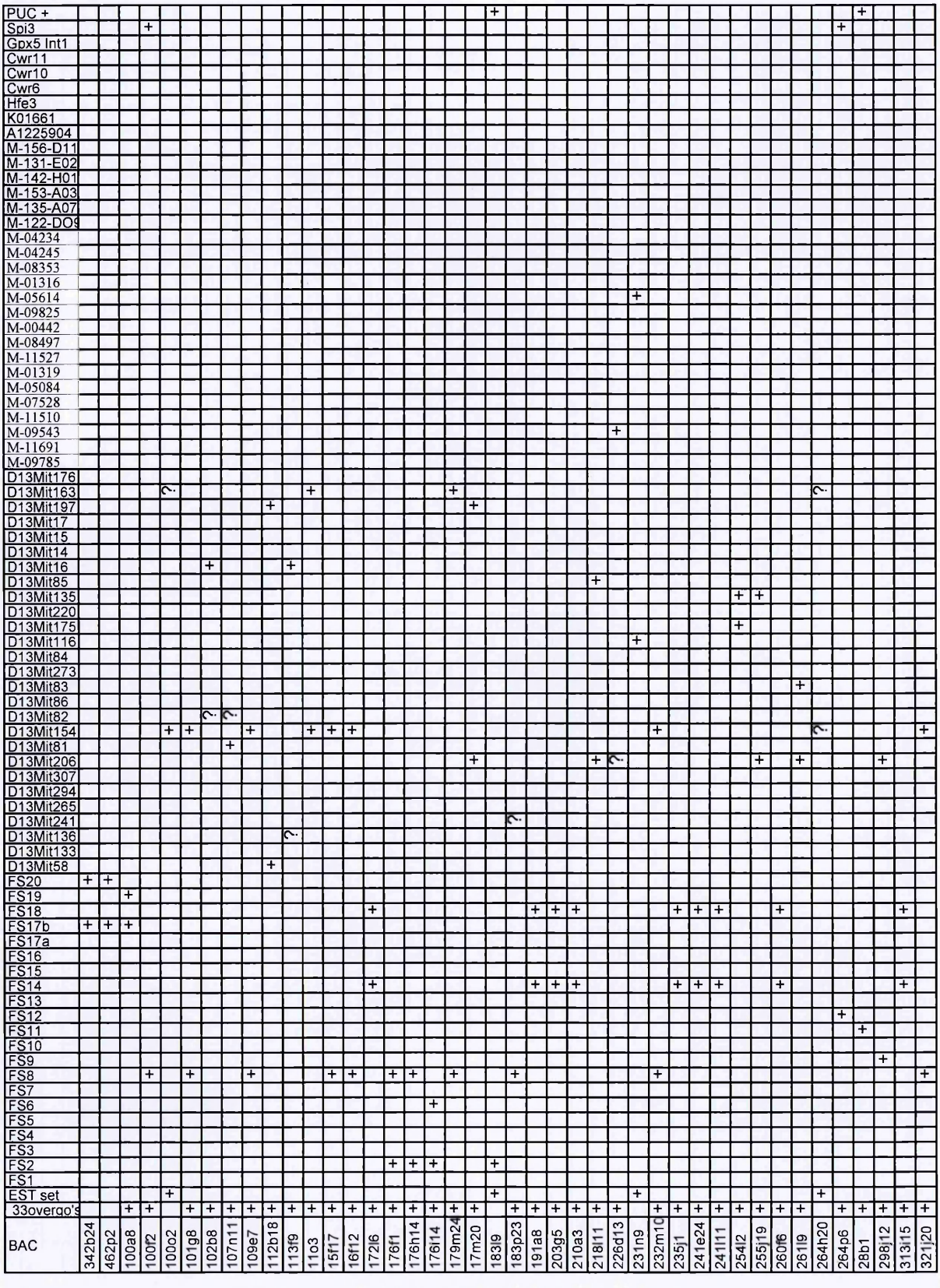

PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-DO9
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
452c7
452p15
458k24
459g14
45i10
462e4
462f13
462g4
462h7
462m19
463d3
466m18
467g5
467k24
468b16
472o12
473i17
475c6
475i22
47a2
480a24
49f22
53b19
53b23
55d12
56d9
57j11
58j15
59c9
5c10
62e12
63j5
63p5
64b22
65a14
65b5
67n14
68h13
71p15
73g8
74n1
78j10

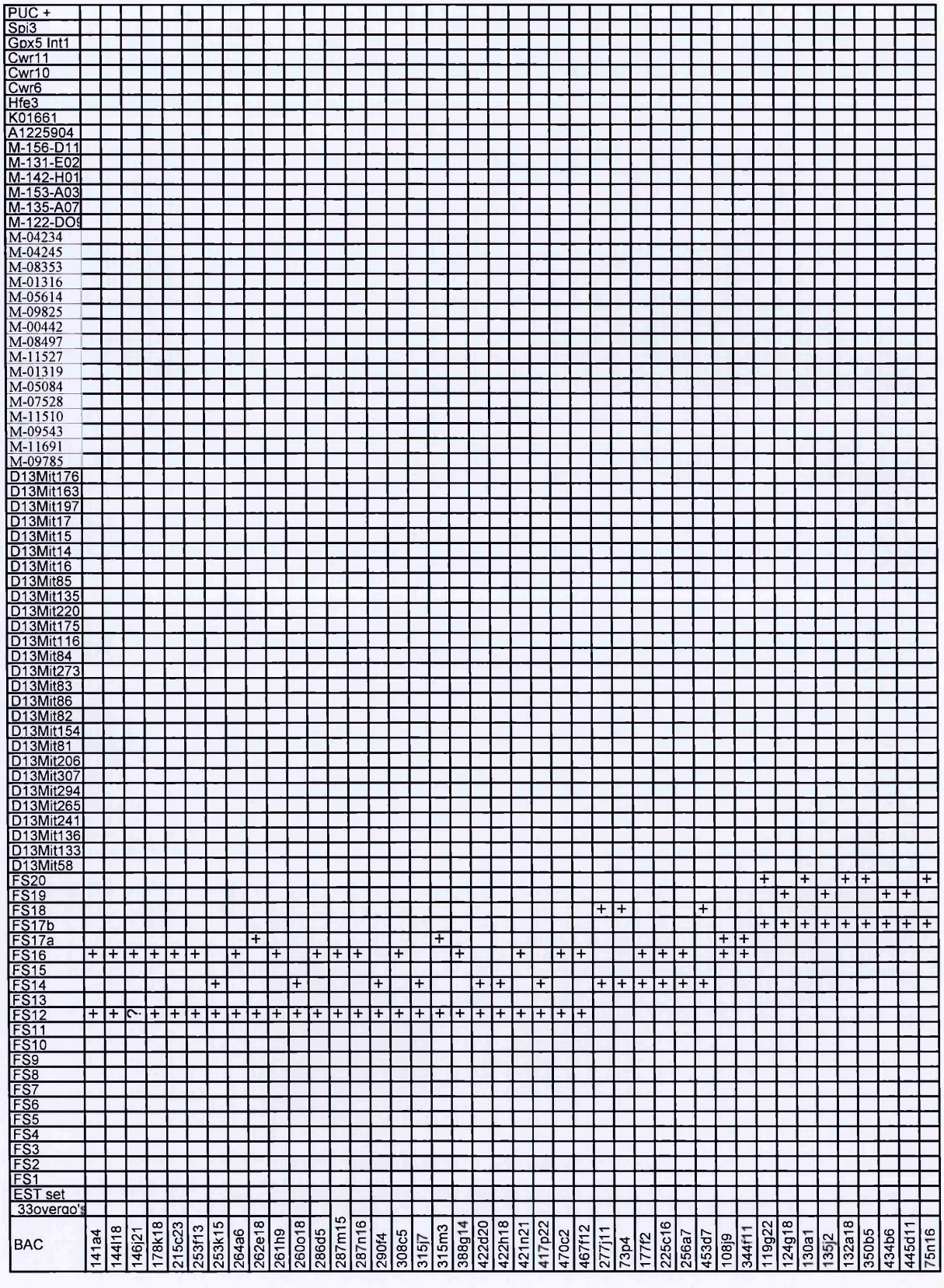

PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-DO9
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
205e2
218a9
224l13
267g7
293b7
298e22
301m19
312k10
307n3
332f3
346f12
386f11
421c4
420e9
423f12
437i5
451h18
461e14
463j22
463k15
462m4
463n15
477b7
22a1
44i6
53h13
89n22
89n23
112o16
116g18
166e12
198p17
208k23
217o20
234n3
248n12
276o1
331g12
331h12
347k14
354a17
388d12

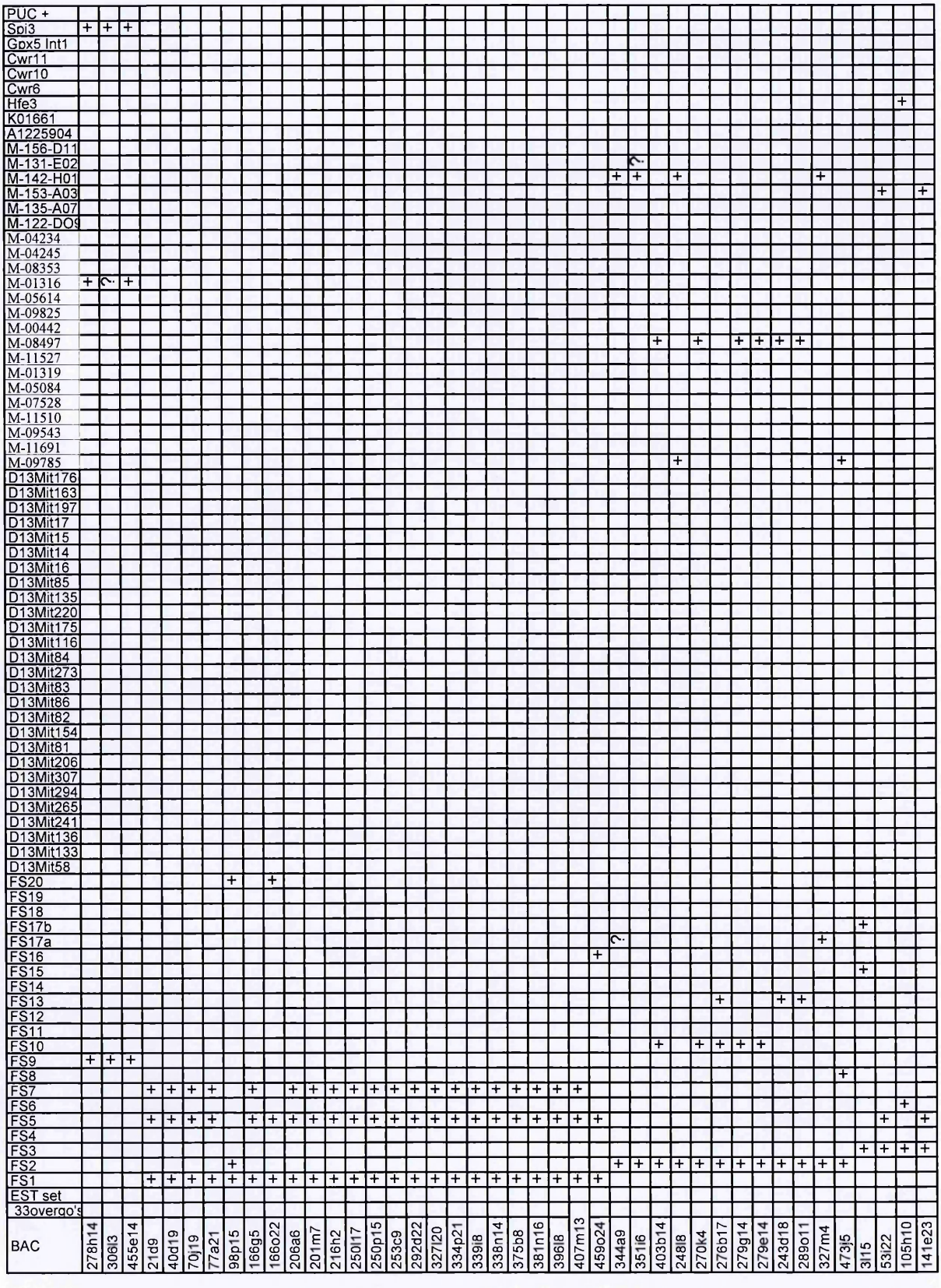

PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-DO9
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
442f5
456g5
464k6
479b14
473c6
476p2
370m11
337a22
384n13
353k5
397m3
417h22
411l16
243b17
311g11
458g11
35e22
57a10
102p15
116p6
134k7
130p18
153a5
286d10
459m24
480a21
85n3
127o17
101k15
31m20
45i4
66b21
79c2
75h1
111o17
126i14
126i15
128j7
126j19
132a9
146k10
154b19

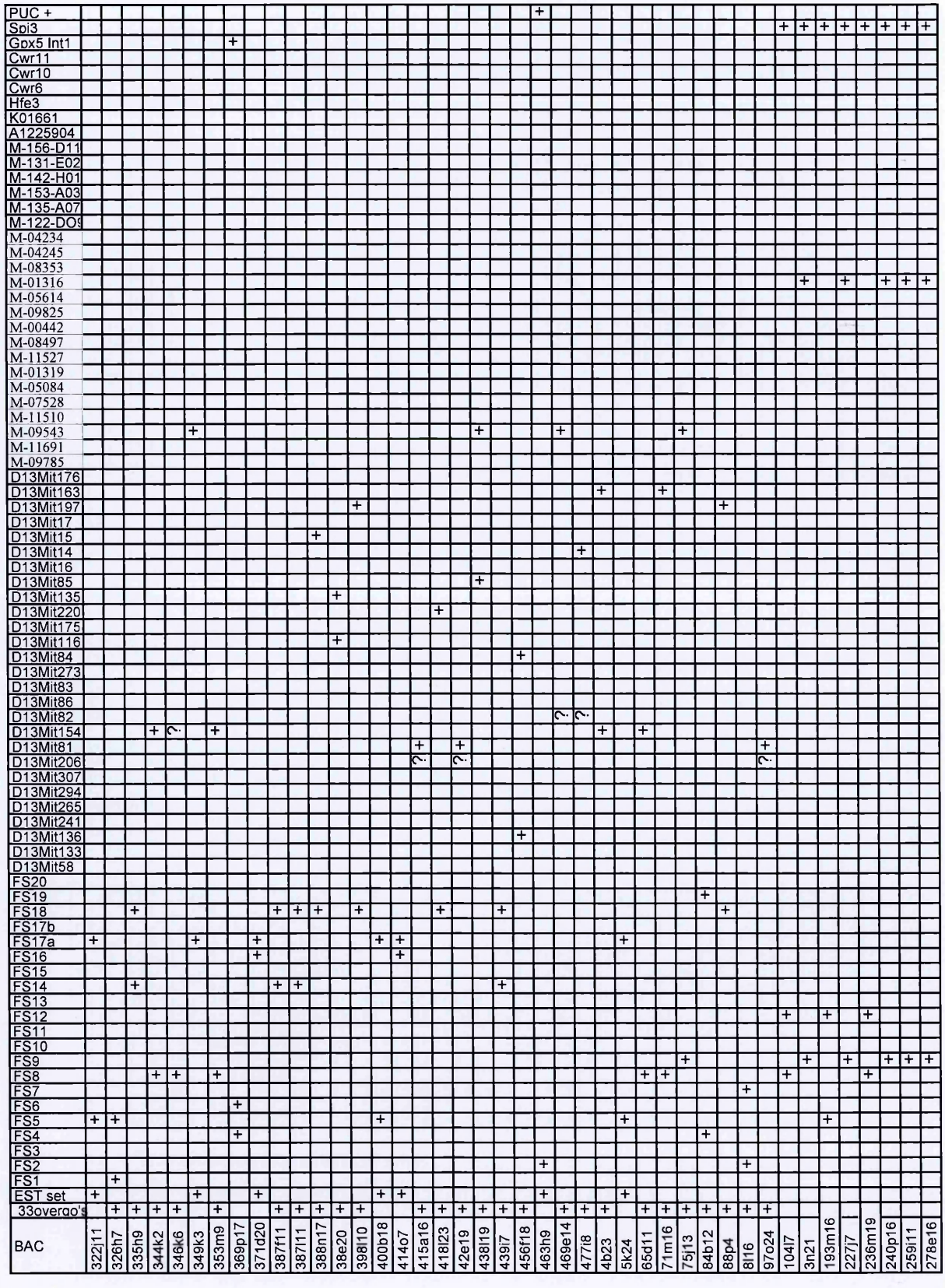

PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-D09
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
180e16
178g14
206g12
204m16
212g3
213g18
217j3
218p21
227f13
259e20
258g15
284h15
290i8
299c17
359a1
370h22
389d23
393a6
408j22
449f24
475d10
58n9
111i18
180m7
240p12
266f10
321j8
340h6
338b5
398g1
458l9
476k14
477f8
28n24
50e15
71f18
71g19
88c1
110d18
122p13
129a7
133e9

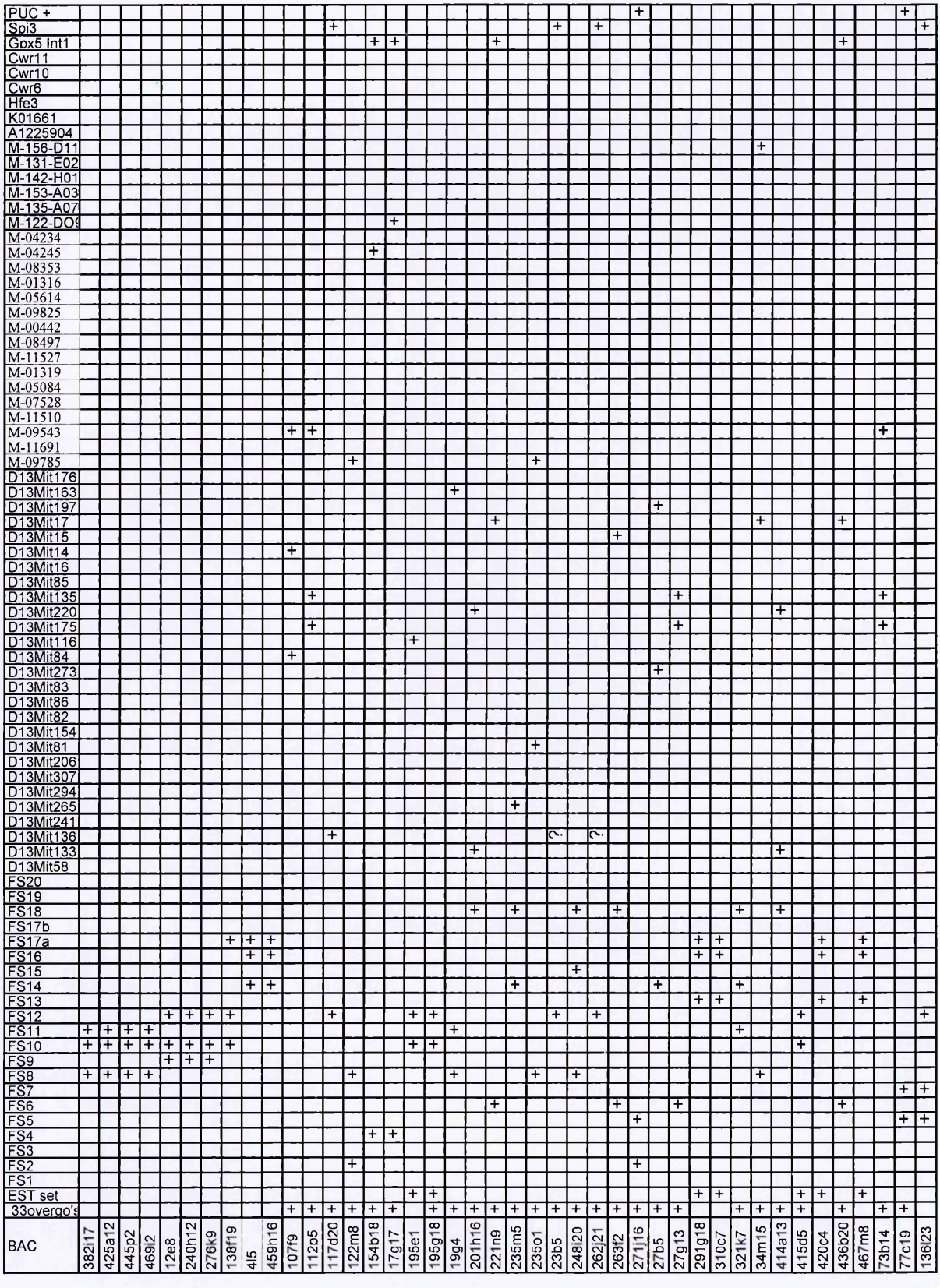

PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-D09
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
386f1
38f15
9o16
128l2
127a11
143b13
254h12
271b10
243n20
279c3
301h12
362n2
424j4
406h19
464j9
220m4
43m4
180j15
180j16
180k16
204c24
226c1
226m1
239b23
236c12
287a24
319e20
461l1
476b6
478h14
478i12
426b24
392e3
39m12
43p12
49f4
58f24
122c17
138b21
138e11
139n19
180e14

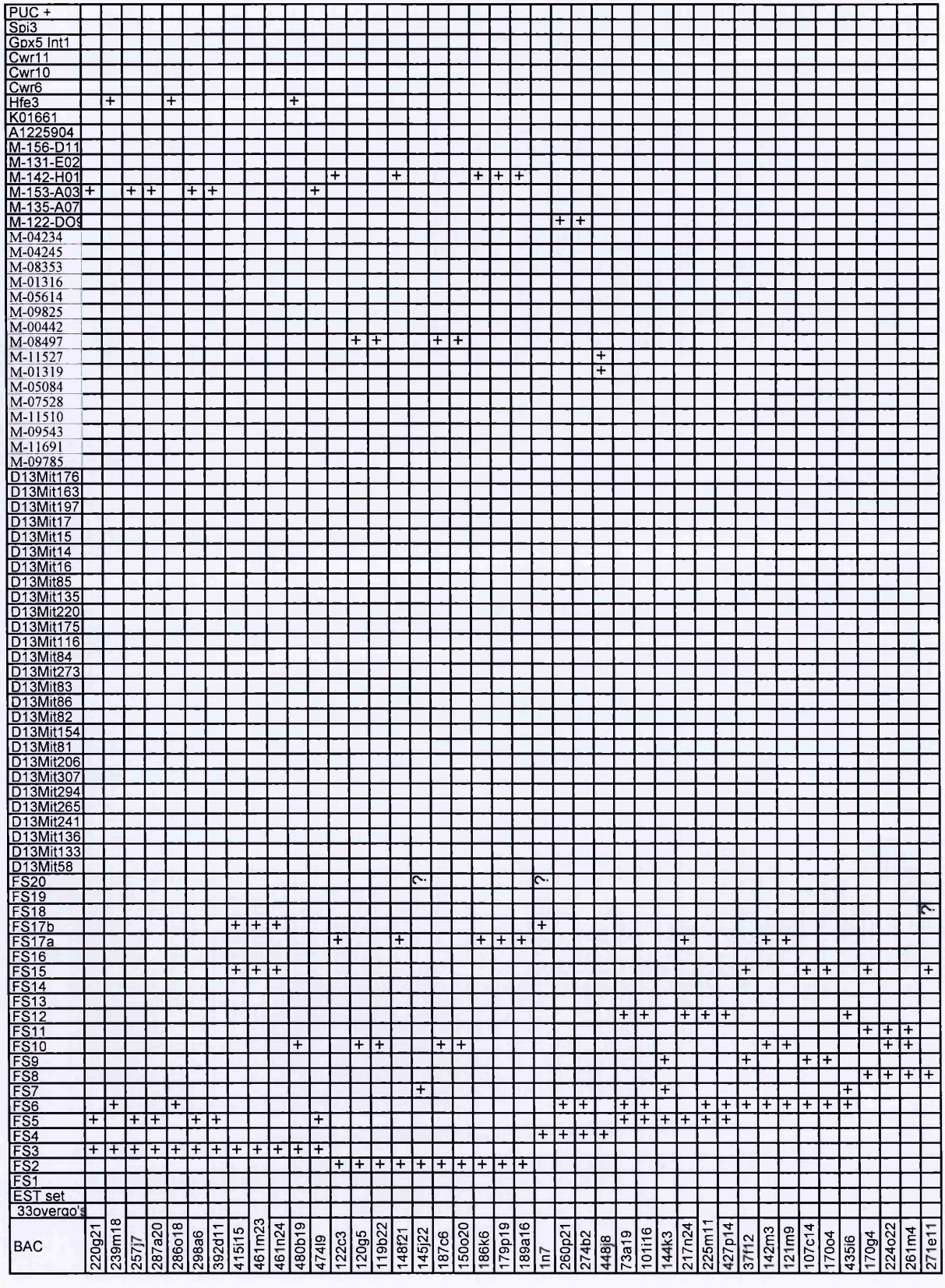

PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-DO9
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
342b24
462p2
100a8
100f2
100o2
101g8
102b8
107n11
109e7
112b18
113f9
11o3
15f17
16f12
172l6
176f1
176h14
176l14
179m24
17m20
183l9
183p23
191a8
203g5
210a3
218l11
226d13
231n9
232m10
235j1
241e24
241l11
254l2
255j19
260f6
261l9
264h20
264p6
28b1
298j12
313i15
321j20

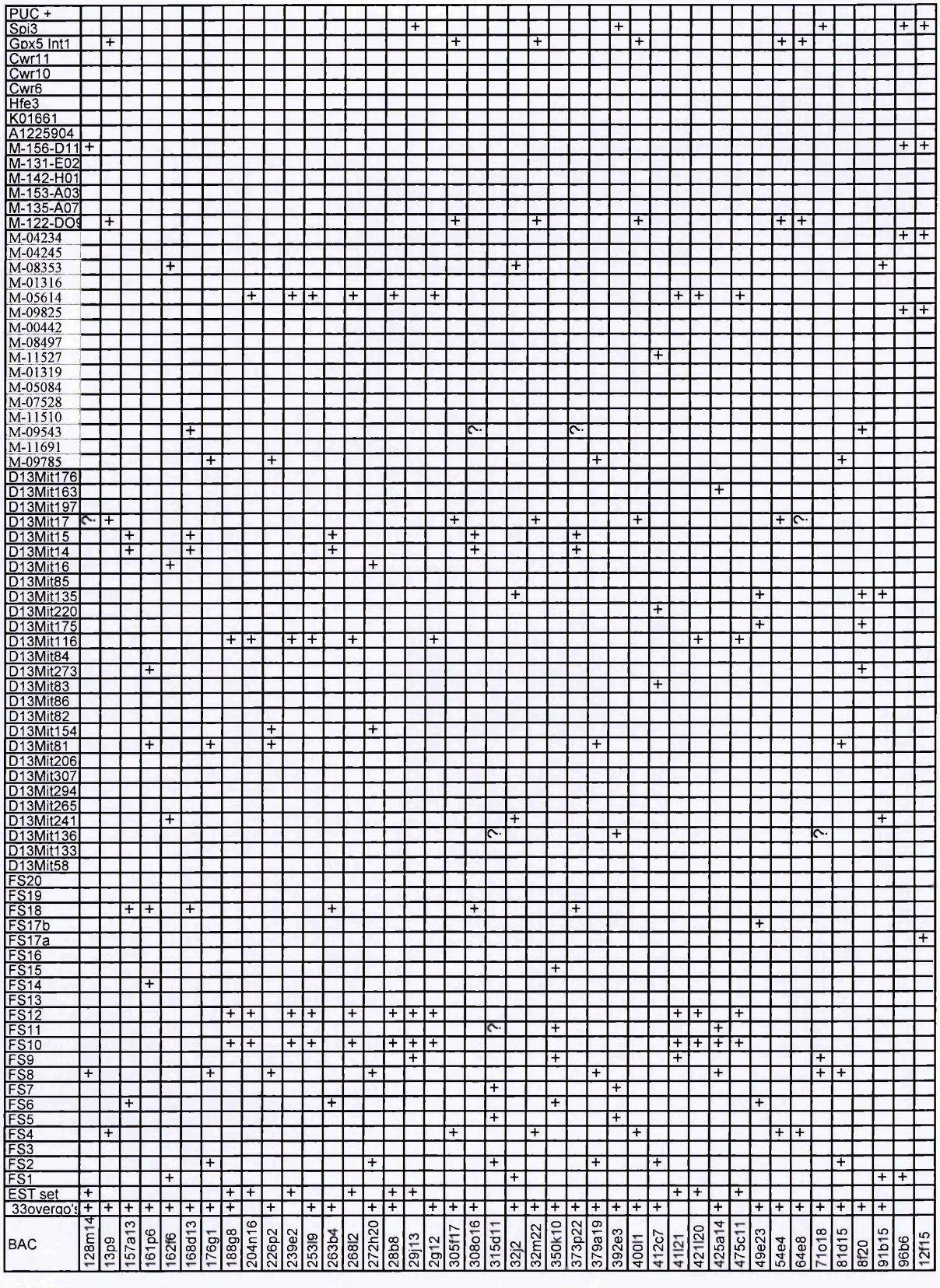

PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-DO9
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
141a4
144l18
146j21
178k18
215c23
253f13
253k15
264a6
262e18
261h9
260o18
286d5
287m15
287n16
290f4
308c5
315j7
315m3
388g14
422d20
422h18
421n21
417p22
470c2
467f12
277j11
73p4
177f2
225c16
256a7
453d7
108j9
344f11
119g22
124g18
130a1
135j2
132a18
350b5
434b6
445d11
75n16

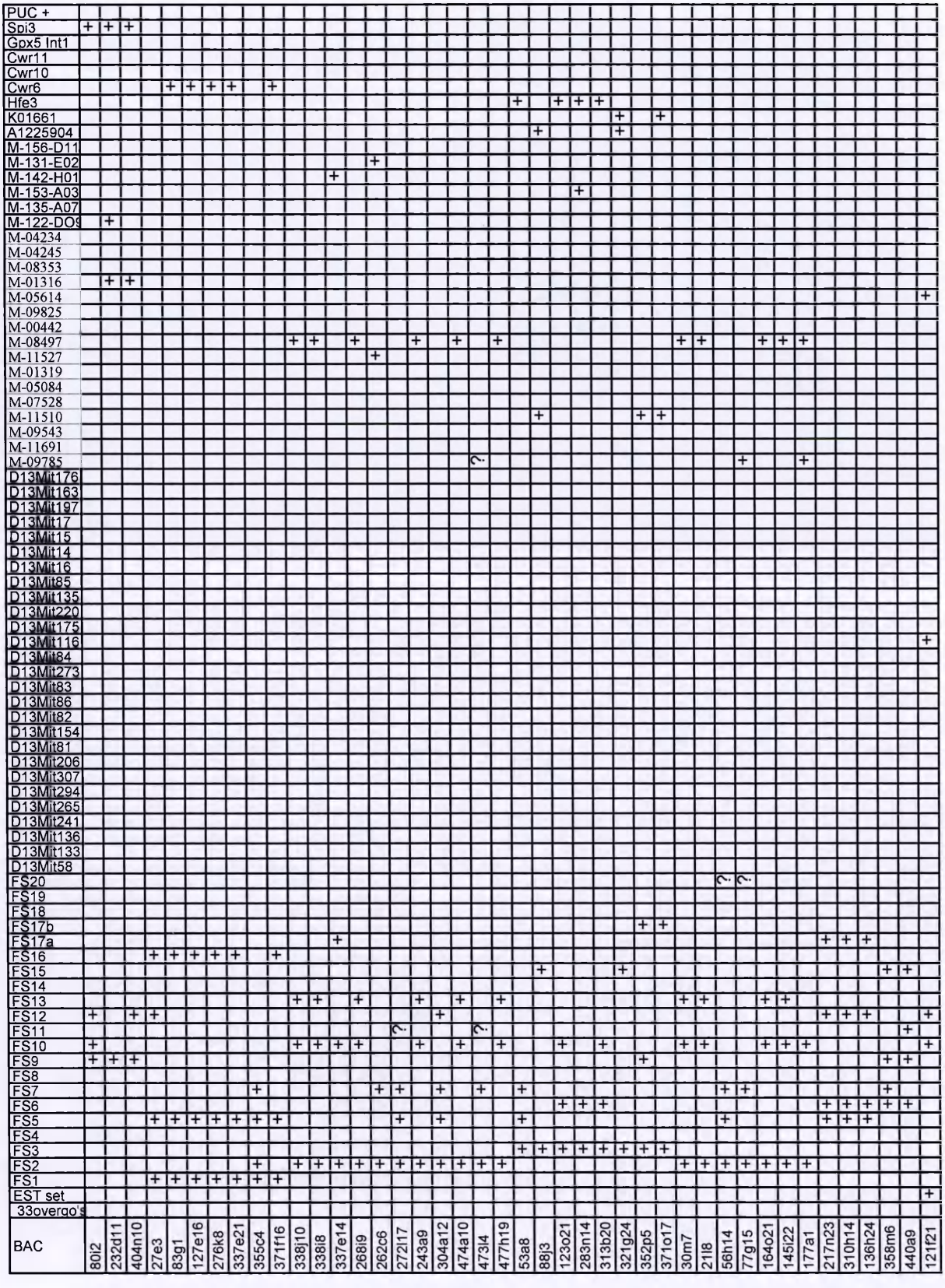

PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-DO9
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
278h14
306l3
455e14
21d9
40d19
70j19
77a21
98p15
166g5
166o22
206a6
201m7
216h2
250l17
250p15
253c9
292d22
327l20
334p21
339i8
338n14
375b8
381n16
396l8
407m13
459o24
344a9
351i6
403b14
248l8
270k4
276b17
279g14
279e14
243d18
289o11
327m4
473j5
3l15
53i22
105h10
141e23

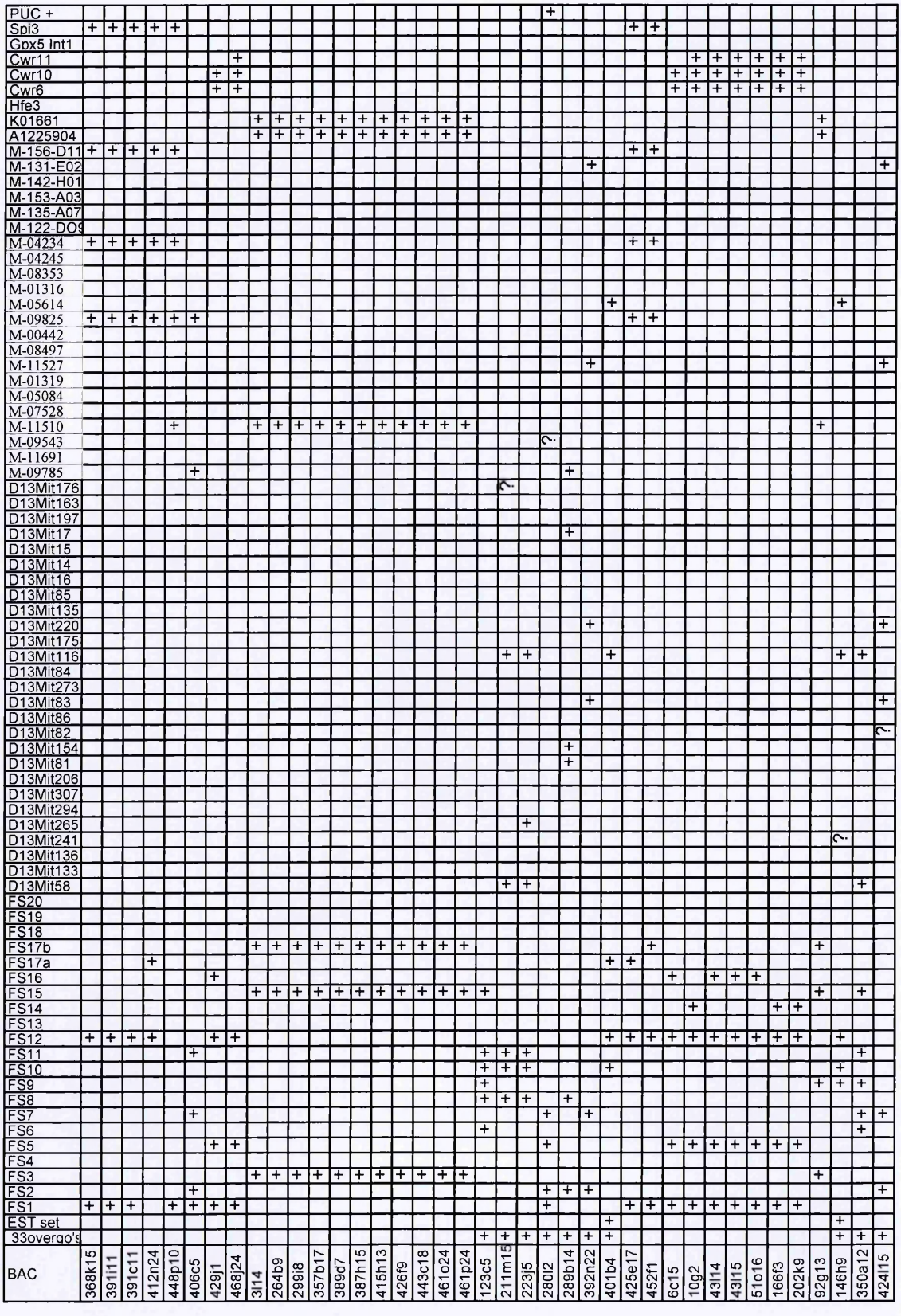

PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-D09
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
322j11
326h7
335h9
344k2
346k6
349k3
353m9
369p17
371d20
387f11
387l11
388n17
38e20
398l10
400b18
414o7
415a16
418l23
42e19
438l19
439i7
456f18
463h9
469e14
477i8
4b23
5k24
65d11
71m16
75j13
84b12
88p4
8l16
97o24
104l7
3n21
193m16
227j7
236m19
240p16
259i11
278e16

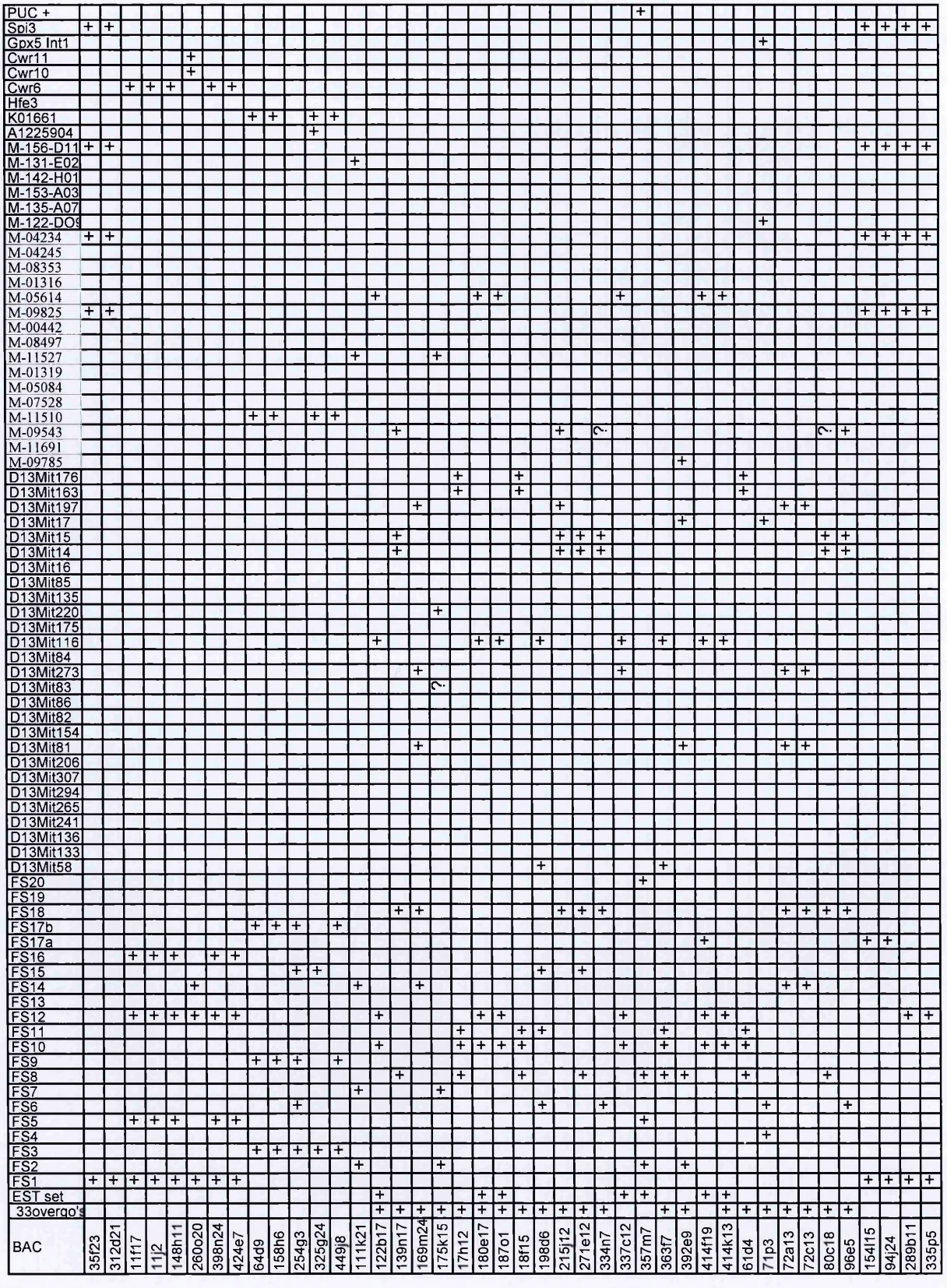

PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-DO9
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
382l17
425a12
445p2
469i2
12e8
240h12
276k9
138f19
4l5
459h16
107f9
112p5
117d20
122m8
154b18
17g17
195e1
195g18
19g4
201h16
221n9
235m5
235o1
23b5
248i20
262j21
263f2
271j16
27b5
27g13
291g18
310c7
321k7
34m15
414a13
415d5
420c4
436b20
467m8
73b14
77c19
136l23


PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-D09
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
220g21
239m18
257j7
287a20
286o18
298a6
392d11
415i15
461m23
461n24
480b19
474l9
122c3
120g5
119b22
148f21
145j22
187c6
150o20
186k6
179p19
189a16
1n7
260p21
274b2
448j8
73a19
101i16
144k3
217n24
225m11
427p14
37f12
142m3
121m9
107c14
170o4
435i6
170g4
224o22
261m4
271e11


PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-DO9
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
128m14
13p9
157a13
161p6
162f6
168d13
176g1
188g8
204n16
226p2
239e2
253l9
263b4
268l2
272h20
28b8
29j13
2g12
305f17
308o16
315d11
32j2
32m22
350K10
373p22
379a19
392e3
400l1
412c7
41l21
421l20
425a14
475c11
49e23
54e4
64e8
71o18
81d15
8lf20
91b15
96b6
12f15


PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-DO9
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
80i2
232d11
404n10
27e3
83g1
127e16
276k8
337e21
355c4
371f16
338j10
338i8
337e14
268i9
262c6
272l17
243a9
304a12
474a10
473l4
477h19
53a8
88j3
123o21
283n14
313b20
321g24
352p5
371o17
30m7
21l8
56h14
77g15
164o21
145i22
177a1
217n23
310h14
136h24
358m6
440a9
121f21


PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-DO9
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
368k15
391i11
391c11
412n24
448p10
406c5
429j1
468j24
3l14
264b9
299i8
357b17
389d7
387h15
415h13
426f9
443c18
461o24
461p24
123c5
211m15
223j5
280l2
289b14
392n22
401b4
425e17
452f1
6c15
10g2
43l14
43l15
51o16
166f3
202k9
92g13
146h9
350a12
424l15


PUC +
Spi3
Gpx5 Int1
Cwr11
Cwr10
Cwr6
Hfe3
K01661
A1225904
M-156-D11
M-131-E02
M-142-H01
M-153-A03
M-135-A07
M-122-D09
M-04234
M-04245
M-08353
M-01316
M-05614
M-09825
M-00442
M-08497
M-11527
M-01319
M-05084
M-07528
M-11510
M-09543
M-11691
M-09785
D13Mit176
D13Mit163
D13Mit197
D13Mit17
D13Mit15
D13Mit14
D13Mit16
D13Mit85
D13Mit135
D13Mit220
D13Mit175
D13Mit116
D13Mit84
D13Mit273
D13Mit83
D13Mit86
D13Mit82
D13Mit154
D13Mit81
D13Mit206
D13Mit307
D13Mit294
D13Mit265
D13Mit241
D13Mit136
D13Mit133
D13Mit58
FS20
FS19
FS18
FS17b
FS17a
FS16
FS15
FS14
FS13
FS12
FS11
FS10
FS9
FS8
FS7
FS6
FS5
FS4
FS3
FS2
FS1
EST set
33overgo's
BAC
35f23
312d21
11f17
11j2
148h11
260o20
398n24
424e7
64d9
158h6
254g3
325g24
449j8
111k21
122b17
139n17
169m24
175k15
17h12
180e17
187o1
18f15
198d6
215j12
271e12
334n7
337c12
357m7
363f7
392e9
414f19
414k13
61d4
71p3
72a13
72c13
80c18
96e5
154l15
94j24
289b11
335p5

Appendix 3.6
Hybridisation Probe Sequences

Mit / Whithead d13 Markers Used for the Design of Overgos	
Mit Marker Used For Overgo Design	**UNISTS Accession No.**
D13Mit136	126959
D13Mit17	116537
D13Mit133	126957
D13Mit15	126972
D13Mit197	127016
D13Mit220	127035
D13Mit265	127082
D13Mit273	127091
D13Mit294	127114
D13Mit306	127123
D13Mit307	127124
D13Mit79	116539
D13Mit84	127182
D13Mit58	127156
D13Mit81	127179
D13Mit175	126990
D13Mit82	127180
D13Mit135	126958
D13Mit176	126991
D13Mit241	127057
D13Mit86	127184
D13Mit85	127183
D13Mit83	127181
D13Mit16	116405
D13Mit206	116540
D13Mit116	126942
D13Mit14	126962
D13Mit154	126976
D13Mit163	126982

Appendix 3.7

Genic Probes Used to Screen the RPCI-23 BAC Library			
Gene	**Accession No.**	**Forward**	**Reverse**
Sox4	X70298	CTCGGGGACTCTAGGTTGG	AGTGTTCTCCGCGTTGTTG
Spi3	U25844	ACAAGGCTGGACAAGATGGA	AACAAGCCTTGCTTGGAAGA
Btn	U67065	TCTCTGAGTGCTTGTGTGGC	CATTGTGGAACAAGGAAGGG
Hfe	AF007558	GGGTATGGCTGTTTGTGACC	ACACAAAGTCAGGGACAGGG
Foxc1	NA	TATCTAGGCCTGCGAGAGACC	CCCAACTCCGAAATGATGG
Olfr11	U28779	CTCTCCATTACTCAGTC	GTGGAAGAACACACTCA
Gpx5 Int1			
Foxf2	Y12294	TCCTGATGGTGGGTCTTCTC	TTGGCTTTCCATATCTTGCC

Appendix 3.8

STS Markers Used to Probe the RPCI-23 BAC Libraries		
Marker	**Forward**	**Reverse**
M-00442	GCTGGTAAGGTCCTCTCTCCT	ATGACACCAGTACCCAGGGA
M-00890	TTGTCCCCATCACTTCATCA	CCAATTCAGTTGAGATCTTTTCC
M-01316	AATTGGTTGCCTCTTGGTTG	CTCCATTTTGGCATCAGCTT
M-01319	CCTGTGCTTTCAAATTGTGC	TGCCTAAAGCATCCTCCACT
M-04234	GTTCTTATGTTTTCCTCCAAGGC	TTTCCAACACCACACCACAT
M-04245	CCATGCCATATACCCTCTGC	GAATTTCAGGGCTTGTGAGC
M-05084	TTCCCCAAGATTTGCTCAAG	TACGCTAGAAAGGAGGGCTG
M-05614	CCCAGCACAAGCTGAGAAG	TTCCTCTTTGTCCTTGTGTCC
M-07528	GCACCAAGGATCTTGTCCTG	TCTGGGAAATAGCAATACTTTTCAT
M-08270	AAGATGCCCTTGAACTTCCC	TGCCTGCCTGTCTACTCTGA
M-08279	TCAAAGTAGCTCTCCCTTTCTCA	TTGAGGGGCAGTAATGTCATC
M-08298	TATCATGGCAAGGAAGGCAT	GGACCCAGCAGCTACAGTTT
M-08353	CACAGCACCAGGTACCACAC	CAAGGAAGAAAACAAACATGTGA
M-08493	TCTTCCTGGCTCTGGGACTA	TACTTGTTTGGGAAGTCGGC
M-08497	TAAAGTTGGGGGTGGTAGGG	CATACATTTTCATCTCTTTCAGACA
M-09543	TGCCACTATGCCTCAAGTTCT	CTGCCCTGCCTCTTGTTAAG
M-09785	CCTCGAGATGAGGCACCA	CATTTGTGGAAGATGCAGAAGA
M-09825	CCAATTTCCCTGAACATAGGTG	CAGCACTCCCTAATCTCTGGA
M-10786	CCACCATACTGTTCCTTCCC	TGGGTGATACCTATGAGGCAC
M-11510	CGCAAGGACAAGACAGGAAT	CCAGGCCACTGATTAGGAAC
M-11527	TTAGAGGCAGGCTCCAGAAA	AAGGCCTGAGTGACTGACAGA
M-11691	GCCCCAAGAGTTCAGTCTGT	CTGCACCTAGGCACATAGCA
Cwr1	AGCACTCGAGGTGGCTTTTA	CAGTTGGGATACTGCCAGGT
Cwr3	CAGTGGGAAGGCTAAAGCAG	GGGAGTATCTTGCCAGGACA
cwr6	GGTCCATCTGTTTAAGGACAGC	ATGTCTTGCTGTTATGGACGG
cwr10	AAGTGTCCCATTGATGCCTC	TCATTTGCCAACAGACCAAA
cwr11	TTCCCACCCTCCTGTAAGC	CTTTCTGAGAAGGCCCGAG
Cwr12	GGCTGTGAAATGCCCTAAAA	TCACGTATCAGAGATGGAAAGG
Cwr13	AGCTCAAGCAGCTTTTATTTGG	CTTGGGCTAAAAAAAAGGGG
Cwr26	TGGTGTTCAGTTACCTTGATGG	GCACCTGGGAGACTTAGAAAA
Cwr27	CCTCCAGAAGCCACAGAAAC	GTTTGTTTCTGCACCCATCA
Cwr30	TTTAAGCCCAGGTCCTCATG	GAGAATGACGGGACCTCAAA
Cwr31	GAGAGACGCTGGTTGTCACA	AGAACCTGGCTAGGACTAGCG
Cwr32	TTTAGCACGAATTGGCTGC	ACTGGGGACTTGAACCCAG

Appendix 3.9

Markers from the Human 6p Region Syntenic to Mouse Del(svea)36H		
Marker	Forward	Reverse
HB001	TAAAAGCCTGTCTGCGGAGT	TGTACCCTTTTGCCATAGCC
HB002	GTTAACAGCCTGGACCAACC	ACATCCACTATTCCCAAGCG
HB003	TGTGGAAGTGAATGAGGAAGG	AAAATGAACGGATGGTCCAC
HB004	TGGAAGAGGGGAAGATGATG	TGCCCTTCTCCATGTACTCC
HB005	TTTGCACCTCAGCTTCAGAG	TCAACAGCATCCTTTGCATC
HB007	TTAATGGGGCTTTGAAGTGC	AATAGCCCACATTTGCTTGC
HB008	TAATAAACTGGGCACGCTCC	CGTTATGTTCATGCCCACAC

Appendix 3.10

EST Probes Used to Screen the RPCI-23 BAC Library			
Assay	Accession No.	Forward	Reverse
AA407021	AA407021	GTTTATTTTTCTTTTAAAAGATGC	TTCAATCATCTGATGAAGAA
AA516957	AA516957	GCTAAGTGCCTAAAACTGCTGC	AAGAGCGACTCCTCATTTCC
AA571384	AA571384	CTCCCCACCTTCCGAC	CACACTGGACCATGACATTATC
AI115341	AI115341	TCCATCATTTTCCAGTCAAACC	AGTGTAAGTTGTGCTTCCTTC
AI225904	AI225904	TGAAAACCTGGACTTTGGATG	TGTCCAGTGTCAAATTCTACGG
R74621	R 74621	AAGAATGCTGAACCAAGAGTTTCG	CCCAAACTCCAAACTCAATTCAAA
K00020	AA407415	AAGGGTCACTTCTTTTCAAA	AATGAGTCAGGGGACAATGG
K00153	AA409771	TGTACTTTACACAAGTGCATT	GGTAATTTTGAAATCAACGA
K00181	AA409971	TTAAACACACAATGTTTATTTTTC	CCTATTGTTTGCATGTTGAT
K00829		CAGCCCTTAATAAACATTCT	TGATAAATGTAACATGCCTG
K00879	C76174	AATACTGACAAGAACAACCTTTC	GTTTCTGTCCAATGTCATACATA
K01193	C77649	TGATACAAACTAAGTTTCATAAAATGA	ATGGAAAGTGCTTTCTATGTAATTA
K01661	C78201	CTAGTTTTACTCCAAAAGAGTAGAATC	GAATATTACATTGTAAAAATAACTGCA
K01931	C79209	GTATTCTACGTGCACTGTAACC	GTATCAAACCAAAATAAAGATGAGTAC
K02035	C79329	CATAGTTTTGAGAATAAGGTATTTGAT	ATTTTACATAGCTATTTTCAGACACTC
K03211		ACAGCATTAATATGTGTTATAAAAGAG	CTACGTTTTATGACTTGTATGGTT
K03265		CATAATTTCAAGTAAAATGAGTATGTG	GAACCAGAGTCCAGTAAC
K03576		GATAATCCACAGGAATGAACT	TGTAATAGAGAATAAGAGTCTGTAGCA
K03922		CAGTTTATTAATTCTTCTGTATTAGAC	CTTATTCCTCAAAGTTTAGACCTT
M122-DO9	AI132259	CAGGACTGGGGAAGGAGATT	TAACCAACATGGGGACCAGT
M131-E02	AA682046	TGCTACTCCCAGCACATCTC	CTACCCACCCCCTCTTTGTA
M135-A07	AI173001	AGGGGCGGTTAGGTTTAGTT	CTTCATTTCTTCGAGCTCCC
M141-B07	AI425970	TGTTTTTAGCTGTGCTTGCC	TTTGTGGGACATTTCCTTCC
M142-H01	AI429535	TGTGGGCCATCTTAGTAGCA	CTCTGAAGCATGGCAAAATC
M151-C02			
M-153-A03	M32460	GGTCTCCTTTCATCACACCG	AAATTCGGTAAATGCCCAAA
M-156-D11	U96700	GTTGCTGTTTTTGCAGAGGA	GGAGCCAGTAGAAGCCAAAG
M161-G05	AF050666	CAACCCACCTCCTTTTCTGT	GCCACCACTACCAAGGATTT
M176-C02			
M199-H07	AI325983	TCTGACCTTGGATTTCGTCA	AAAGAGAAGGGCAACAGTGG
M203-B05	AI325983	CTGGGGCTTCACTTAATGGT	GCATCAGCCATGCTAGAAGA
M203-D06	AI325983	TCAGTCCAGGGTTATGTTTCC	TCTTGACTGGCACCTACTGC
M211-E01	AI325983	AAAACAATGTCCAAGAGGGG	TTGTATGCCTCTGGAAACCA
M212-E04			
M220-E01	AI841875	AAGTTGCTGGAGGCATTCTT	GTTCTAGAGTCTGGGGCAGC
M237-F06	AI648199	GAGTGAAGTGCAACAGAGGC	TTGCGCAGTAGAGTCAGGTC
M239-C02	AV290867	GCCCATCAAAACAACTGCTA	AAAAGAAACCAATTGCACCC
M253-H12	AI385587	AGCAAAATCCATGGCCTATC	ATCTCTGGCTTCGAGCTCAT
M257-E09	AI528624	ACCAAGCCTACAACTTGCTGG	TGATCGTGAACTGGAGCTGC
M265-G02	AI875505	TCACTGGTGAACTCGTAGGC	CTTGGTTGTGTACCCATTGC
M275-F01	AW261723	AGAGTCCTGGCTGAGGAAAA	TCATGGGAACATCAAGAGGA
M280-D05	AW492955	GCACTCATCTACCCACAGGA	TGATCAAATTATGGTAGGTGGC

M284-F11	AW538719	TTCCTCCTCTTCCTGTTTGG	GAGCGGTGTCATTCTTCTGA
M285-C08	AW540195	AGCATGACTTTTGCATGGAG	TAGTTTGGGAACAGGTGCAG
M289-B08	AW546347	CATGTACTGCACGTGCTCAG	TACGAAGCGCTGAAAGAGAA
M301-H08	AF011385	TGCCTACGCGTTGATATATTCTT	AACATACAAATGCATTCAAAAGCTA

BAC End Sequence Markers Used to probe the RPCI-23 BAC Libraries

Sequence ID	Forward	Reverse
112o16t7	AAATAGAGCGAGCCCTTAGC	CAATGGCAATGGCTATGATG
11F17sp6	AAAACCGTGTGGTCTGGAAG	ACCACCACTAGAGTGCAGCC
122c3sp6	CAAAAGAAAAGATTTTGGGCTTAC	AGATTCTCTTCTCCCCACACC
12F15t7	TTGCCTCTAGACCACCCTTC	TGCATGCGAAATAGATGTGG
148F21t7	AACAACACTGCTTTTCATTACCC	TCTTCTCGACTGTACCTCTTGG
16F12sp6	CAAGTCTCGTTTCTGGAGGG	TACCTGGTGATCCAAAAGGG
16F12t7	CGGGATATAGTAGAGCCTCGG	TCACACAGTCCACTCATCCC
18F15 t7	GCCCCAGTGTCTTGTGTTAG	AAGAGGGGCAGGCACTAAAG
18F15sp6	AGGAAAGCCCTGGGTATTTG	TCTCTGGGTGTTTCTGGGAG
241e24sp6	TCAGCTTCCCCAAAATCATC	CAGAGCCCCACTCTGATAGG
254H12sp6	TGCCTTGATATTTGGATTTGC	CAGACTATGCAAACACATGGAG
266F10t7	GCCTGTGAGATGCTGTGTTG	AAGTTCCTTCCAGCAGGTTG
271B10sp6	TCCCAAATCAGGAAAACTGC	CCCTTTCTTCACCTGAGAGG
278e16sp6	ACACAAAGATAACCGCAGGG	CAAAGTTCTCCAGCTCAGGG
278e16t7	TTCAGAGACACACAGGTGGC	ACACTCCGGGAACTTCAATG
27E3sp6	TTCCCATCCACAGAGGCTAC	TGAAACTCTGGGAGCAGATG
290i8t7	TGTAGTGACACCGCCAGAAG	TGGCAAGAGTCACCTTCTCC
298j12sp6	CCTGTAGCGCTGTGAGAGTG	CTGTCCTACTCCCCTCCCTC
310c7sp6	TCTTCCCCTTCCCAGATAGC	CTTGGTTAGGCGAGGAAATG
310c7t7	AATGGTGGGTGTCAGAAAGC	GGGCTGTACACACACTGAGC
312d21sp6	AGATCCTGCACTCCCAGTTG	CCTTTGAGATGTGGCAATTATG
326c24sp6	AAGCCATCTGGACACAAACC	GGAGCTCCATTCTCATACCG
335h9t7	AGTGAGGTGGGGAGACATTG	CCTCAGTACTGGCTGAAGGG
349k3sp6	GTCCTTCCTTCCCAGTGTTG	CAAATGACCCTTGGAAGTCG
34M15sp6	CGCTACTTCGCTTTTATCCG	TGGAGTCTGGGAAGGAACAG
34M15t7	TGGCAAGCACTTAACCACTG	ACTCCATTTGCCTTTTGTGG
392L15sp6	AGGGACTGGCCTTCTTTCTC	GCAGAAAACTGGGGTAGCAG
39M12sp6	TTTAACAGGATCAGGCCACC	TCAGCTGTTTGACCTGCTTG
404N10t7	ATTGAGGGACATCTGCGTTC	TCCATCCTTCCTGTTTCCAG
408j22t7	TCACATTAATGCCTGCCATC	TGAGTCGCCTGTGTTGACTC
414k13sp6	CAAATGGCCAGGAAGTATGG	CCCAAGAAGAAGACAGCCAG
439F14sp6	ACATGAGCTGTTGGAGGAGC	CAAACCTCCCACCACTGAAC
467F12t7	AGGTGGGGCATCCAAAGTAG	AGACTTCCAACCCAGTGCAG
468M24sp6	TCCATGATGTTTGGGTGATG	TTGAAGCGTCTTCCACACTG

509H17t7	GCTCCAACTTCCTGCTCAAG	AGGAGAGTACTCAGCCCATGTC
51a11sp6	ATTTCGGCAACATTAGCTCG	AGGAACCTAGGCTCTGGGAG
584b16t7	GCCTACTTCTGTCACATGGC	CAGAAACATGGACATGGTGG
71F18t7	TGCCTCATTTGCTGTCTGTC	ATTGAGTACGTGGGTCCTGG
72c13t7	GGCTGCAGTATTTAGTTTGGC	TCCCCATCCTTTGTCTTCAG
80c18sp6	GAAAAGTGCTAGGTGCTCGC	AAAGTGAAGGGGAAAATGAATG
80c18t7	TAGTCTGTGGGAGGGGTTTG	GGCCATGTGAGTGTTGTGAG
87c10sp6	TGCTGCTAAGACATCCATGC	ACCCAAAACCAACCAACAAG
92g13sp6	CCCCAGTGATACCCAGTGAC	TTCCATGGGCGTTAAGTTTC
92g13t7	TGCACAGAAGCCAGAGAAAG	GTCTGGGGCTTTTGAGACTG
94b10t7	TCTGTGCTAGTGTTGGCTGG	GCGAAGCTCCTAGATGGTTG

Appendix 3.11

PAC End Sequence Markers Used to probe the RPCI-23 BAC Libraries		
Sequence ID	**Forward**	**Reverse**
343e12sp6	CAGCACCCAACTACCATAGG	AAACCATGAAACCAAAAGAGATG
366h7t7	CCGGTAGCAGCATAAACCTC	GCATGTGGATGTGGTCTGTC
423H18sp6	AATCACAATGACGCCTCCTC	GAATCATGCTGGGATGGAAG
490K5sp6	TTGGCTACATTAGTTTTCATGAGG	GTGGAAAGTTGTAGCCCTGC
525F18t7	ACTTCTTCAGGCCTGTGTGC	CTTCCCAAATGGTGTGTGTG

Appendix 3.12

Primers Used to Screen For SNPs within the Del36 Region		
Primer ID	**Forward Primer**	**Reverse Primer**
c13snp1	TGGTGACCCATTGATTTTCC	AGACCACGAGAGGATTGAGC
c13snp2	TTCCTTCAAACGTGGTTTCC	AAGAACACCCCAACCCTTTC
c13snp3	GCAGGCTGAATTTCCCTTATC	GACTCCCCGAGAAGAAGAGG
c13snp4	TAGCAATGCACATGGCAAC	CAGCAGGTCACATGATGAGG
c13snp5	GTGCCTTTCCCCTTCTTAGG	ATTGCCTTGCTGGTGGTATC
c13snp6	GCAACCTCAAAGAGGCAGAG	TTTTAAGCCTGCGTCTACAGC
c13snp7	ATGCAACAGAGCCACATAGC	GGCTCTGACTGATGGTTTCC
c13snp8	TTAAGGCCTTCCATGCTTTG	TCATGTTCACTTGCCCAAAC
c13snp9	ACACCAAAAGGAGCATCAGG	GCCAAGCTGTCTATCCAAGG
c13snp10	GGGCTTCAGCTCTACAATGC	GGTGTCAGGGACAGCAGAAC
c13snp11	GAAAGCTCGTCCTTGCAAAC	ACAAGAGCAGAGGTGGCATC
c13snp12	GGAACAGCTACCCACTGAGC	TGGTGAAAAGCTCCTGGTTG
snp1b-R	TGGTGACCCATTGATTTTCC	TCATGCAGCATCGTCTAAGC
snp2b-R	TTCCTTCAAACGTGGTTTCC	AAAATTCTGCGCAACCTGAG
snp3b-R	GCAGGCTGAATTTCCCTTATC	GGAAGAGTAAGTCCCGCCTAAG
snp4b-R	TAGCAATGCACATGGCAAC	GGGGATTCTCCGTCAAGTTC
snp5b-R	GTGCCTTTCCCCTTCTTAGG	GTGACCTTCTCCCTGGTTCC
snp6b-R	GCAACCTCAAAGAGGCAGAG	TGTAAGGTAGGAGAGAGTGGTAAGC
snp7b-R	ATGCAACAGAGCCACATAGC	TGGAAACTCACCAAAACCTTTC
snp8b-R	TTAAGGCCTTCCATGCTTTG	TCTGAGCGCAGCTTTCTATG
snp9b-R	ACACCAAAAGGAGCATCAGG	CTGCAAAGTGATTGCTCTTCC
snp10b-R	GGGCTTCAGCTCTACAATGC	GTCTCGATGTGGCTCAGGAC
snp11b-R	GAAAGCTCGTCCTTGCAAAC	TTTGAGCAGGGACTCTCAGG
snp12b-R	GGAACAGCTACCCACTGAGC	AATAAAGGCCATCCAAGTCG
snp1c-L	CATCTCCCACAGTTCCAGTG	AGACCACGAGAGGATTGAGC
snp2c-L	GAAGAGACTGGCCTCTGAGC	AAGAACACCCCAACCCTTTC
snp3c-L	AACCCCAAAATGGTTGGAAG	GACTCCCCGAGAAGAAGAGG
snp4c-L	ATCTGGCCTGAATTGGTTTG	CAGCAGGTCACATGATGAGG
snp5c-L	GAGGCACAAGAGCAGAATGG	ATTGCCTTGCTGGTGGTATC
snp6c-L	GCCTGTGACCAGCAAGATATG	TTTTAAGCCTGCGTCTACAGC
snp7c-L	AAGCACAAATGCCACATACAAC	GGCTCTGACTGATGGTTTCC
snp8c-L	TCGCCTTGCTGAACAGAAC	TCATGTTCACTTGCCCAAAC
snp9c-L	GACACCAGGTTCTAGCTGACG	GCCAAGCTGTCTATCCAAGG
snp10c-L	AAAGCCAGCCTAGGTCACAC	GGTGTCAGGGACAGCAGAAC
snp11c-L	TCAAAGGTGCTTTTGCAATC	ACAAGAGCAGAGGTGGCATC
snp12c-L	GTGGCATTTGCCTTCTTTTC	TGGTGAAAAGCTCCTGGTTG

Appendix 3.13

Primers Used for Genotyping			
Primer ID	**Forward Primer**	**Reverse Primer**	**Forward primer Fluorescent Label**
Foxq1_F2&R5	GAGATCAACGAGTACCTCATGGG	CGAAGGAGCTGGAGAACTTG	
D13Mit115(+FAM)	TGGTGAAGTGTTTGGAAAAGG	TTTAACCCATTGATCTACTTCAAGG	FAM
d13Mit17(+FAM)	CACCCCCAAGTTCTCTTGAA	CCCACATACACATGTGCACA	FAM
DNR059(+ROX)	GTTCTCACGGCCATTTCAG	CCATGATGTGAGCAAGAAACC	ROX
DNR059(alt1)		TTTTGCCCTATCCAGACCAG	
DNR059(alt2)		CGGAAAAGGAAAAGAATAAAAGG	
TetNR5(+FAM)	TCTCCTGCATGCTCTGTGTC	AACAGCCTGGTCTATAATGCAAG	FAM
TetNR10(+FAM)	GCTGATGCTAGGGCTTTCTC	CCCTAGGGCACATTTCACTG	FAM
DNR037(+ROX)	AACTGCTGGGGAAAGGTAGG	CACGGGGACCATTACTTCC	ROX
DNR024(+HEX)	AGCCCACTTCAACTCCTCAG	CCCTCAGGGGTACCTCAATC	HEX
DNR024(alt1)		TCACTCAACACTTACATCATAGGC	
DNR024(alt2)		GGGGTACCTCAATCACTCAAC	
DNR003(+HEX)	AAGTCAAACATGGCTGCTAGTG	GGAGAAAACAGACTTCATTCAGG	HEX
DNR003(alt1)		ATAAGGCATTGGCGAGAATC	
DNR003(alt2)		TAAAGGGCATTTCCTTGGTC	
DNR016(+ROX)	GGAGAATTGTGAGGGAATGC	CAGCCATCACATACCACACC	ROX
DNR016(alt1)		TCCTAAGCCACACCCAAGAG	
DNR016(alt2)		ATCCCTGGAACCTCAAATCC	
d13Mit136(+FAM)	TTTTATCTATTGAGTAGATTCATGGTG	TATGCCTGGAGGAAAACAGG	FAM
dnr123(+FAM)	GAGGAGTGTATGGCCTGATTTC	ACAGGAGCAGCTTGTCATTG	FAM
DNR123 Ri		CGGCAGGATGCCTATTATATTG	
DNR123 Rii		TGATAAGTAGTTCCAGGCCTTAGC	
DNR123 Riii		CCACAGATCATCTAGCTTGTGC	
d13Mit15(+FAM)	AGGAACAGCAAGCTCTAAGGG	GGCCTCCCAAGATATCATCA	FAM
d13Mit15 Ri		CCGTTTAGGTTTTTAAGTTACAGG	
d13Mit133(+HEX)	TAGACACTTAATTCTGTGATGAAATGG	AGCAAAAGCCCCAGTTAGTG	HEX
dnr48(+HEX)	CATGCAGTCAAAACATGCTG	CCCTACACACAAGGTTCAGG	HEX
DNR048 Ri		GCAGATTCAGGTGTCTGAGG	
DNR048 Rii		GGGTTACTTGTATCACCCTACCC	
DNR048 Riii		GAGGGTTATTGAGATCTTGATGG	
dnr76(+FAM)	CATTAAGCTCCACCCATTCC	ACAGGAGCAGCTTGTCATTG	FAM
DNR076 Ri		ACAAAGAAGTGCTCCCAACC	
DNR076 Rii		CCCAGCTTAGCAAACCAATAC	
DNR076 Riii		ACAGGTGTGCATCAACAGAAC	
dnr101(+HEX)	GCAATTAGCAGGAAGGTGATG	CAAAACACCCGAAGGAACAG	HEX
DNR101 Ri		TTTCTTCCGCTTCCATTAGG	
DNR101 Rii		CTTCCAACTGGGCCAAAAC	

DNR101 Riii		TCCATTAGGCAGATCAAATGTC	
d13Mit198(+ROX)	TTCATGAGTCCCTAACCAATAGTC	TCTTAAATAGTCCATGCATTGGC	ROX
DNR111(+ROX)	ATTGGCACTTGCCTGGTAAC	TGACTGTCATGCCTACACCAG	ROX
DNR111 Ri		AACAGAGACCCTTTGTGAGAGC	
DNR111 Rii		GACTCCTGAGTGGTTGTGAGC	
DNR111 Riii		AAGCACAGCAGGAAAACTAACAG	
DNR201(+ROX)	ACACACACGGCAAGTGCTC	TTCATGGCTCCACTGTTCTG	ROX

Appendices for Chapter 5

Appendix 5.1

Table of Results from all urinalysis tests

Urinalysis												
ID	Genotype	Sex	Date	Bilirubin	Urobilinogen	Ketones	Ascorbic acid	Glucose	Protein	Blood	pH	Nitrite
8.1g	+/-	M	18/06/04	neg	norm	neg/+	++	neg	+	neg	6	n
8.1h	+/-	M	18/06/04	neg	norm	neg/+	++	neg	+	neg	6	n
8.1j	+/-	M	18/06/04	neg	norm	neg/+	++	neg	+/++	neg	6/7	n
8.1l	+/-	M	18/06/04	neg	norm	neg/+	++	neg	+/++	neg	6	n
10.1d	+/-	M	18/06/04	neg	norm	neg/+	++	neg	+	neg	6	n
10.1e	+/-	M	18/06/04	neg	norm	neg/+	++	neg	+	neg	6	n
6.1f	+/+	M	18/06/04	neg	norm	neg/+	+/++	neg	Neg/+	neg	6	n
6.1h	+/+	M	18/06/04	neg	norm	neg/+	++	neg	Neg/+	neg	6	n
7.1e	+/+	M	18/06/04	neg	norm	neg/+	+/++	neg	Neg/+	neg	6	n
8.1k	+/+	M	18/06/04	neg	norm	neg/+	+/++	neg	+/++	neg	6	n
9.1g	+/+	M	18/06/04	neg	norm	neg/+	++	neg	Neg/+	neg	6	n
9.1h	+/+	M	18/06/04	neg	norm	neg/+	++	neg	+	neg	6	n
9.1i	+/+	M	18/06/04	neg	norm	neg/+	+/++	neg	Neg/+	neg	6	n
11.1e	+/+	M	18/06/04	neg	norm	Neg	++	neg	+	neg	6	n
6.1d	+/-	F	18/06/04	neg	norm	neg/+	++	neg	neg	neg	6	neg
8.1c	+/-	F	18/06/04	neg	norm	neg/+	+/++	neg	+/++	neg	6	n
8.1e	+/-	F	18/06/04	neg	norm	neg/+	+/++	neg	+	neg	6	n
9.1d	+/-	F	18/06/04	neg	norm	neg/+	+/++	neg	+	neg	6	n
9.1e	+/-	F	18/06/04	neg	norm	neg/+	+/++	neg	+	neg	5/6	n
6.1e	+/+	F	18/06/04	neg	norm	neg/+	++	neg	+	neg	6	neg
10.1b	+/+	F	18/06/04	neg	norm	neg/+	++	neg	+	neg	6	n
11.1a	+/+	F	18/06/04	neg	norm	neg/+	++	neg	+	neg	5	n
11.1c	+/+	F	18/06/04	neg	norm	neg/+	++	neg	+	neg	6	n
9.1a	+/+	F	18/06/04	neg	norm	neg/+	+/++	neg	+	neg	5/6	n
8.1a	+/+	F	18/06/04	neg	norm	neg/+	+/++	neg	+	neg	6	n

Appendix 5.2

R Code for Fisher Exact Test.

Fisher Exact Test for Urine Ascorbic Acid in Males

```
> males_ascorbicacid <- matrix(c(0, 6, 4, 2), nr=2,
+ dimnames=list(c("+/++", "++"), c("muts", "wt")))
> fisher.test(males_ascorbicacid)
```

Fisher Exact Test for Urine Ascorbic Acid in Females

```
> females_ascorbicacid <- matrix(c(4, 1, 2, 4), nr=2,
+ dimnames=list(c("+/++", "++"), c("muts", "wt")))
> fisher.test(females_ascorbicacid)
```

Fisher Exact Test for Urine Protein in Males

```
> males_protein <- matrix(c(0, 4, 2, 5, 2, 1), nr=3,
+ dimnames=list(c("neg/+", "+", "+/++"), c("muts", "wt")))
> fisher.test(males_protein)
```

Fisher Exact Test for Urine Protein in Females

```
> females_protein <- matrix(c(1, 3, 1, 0, 6, 0), nr=3,
+ dimnames=list(c("neg", "+", "+/++"), c("muts", "wt")))
> fisher.test(females_protein)
```

Appendix 5.3

Frequency of Animals Tested for Blood Metabolites.

The number of animals used in each of the Blood biochemistry tests. Grey boxes indicate that fewer than three animals were tested in one of the genotype groups.

Number of Animals used for each Metabolism Test.								
Test	Test 1				Test 2			
	Males +/+	Males +/m	Females +/+	Females +/m	Males +/+	Males +/m	Females +/+	Females +/m
Sodium	8	5	7	4	8	4	6	4
Potassium	8	5	7	4	8	4	6	4
Chloride	8	5	7	4	8	4	6	4
Urea	8	3	7	4	8	4	6	4
Creatinine	8	3	4	4	6	4	6	4
Tot. Calcium	8	3	3	2	5	4	5	4
Phos	3	0	0	1	5	3	3	2
ALP	5	3	1	1	4	4	4	2
ALT	5	3	2	2	3	4	5	4
AST	1	1	0	1	3	4	3	2
Tot. Protein	5	2	0	1	3	4	3	2
Albumin	3	1	0	1	3	4	3	2
Tot. Cholesterol	0	0	0	0	3	4	3	1
Triglycerides	0	0	0	0	2	3	1	1
Glucose	8	3	6	4	7	4	6	4

Appendix 5.4
Blood Metabolism Data.

Metabolic bleed: Test 1															
ID	Sex	Genotype	Sodium (mmol/l)	Potassium (mmol/l)	Chloride (mmol/l)	Urea (mmol/l)	Creatinine (mol/l)	Calcium (mmol/l)	Phos (mmol/l)	ALP (U/l)	ALT (U/l)	AST (U/l)	Tot. Protein (g/l)	Albumin (g/l)	Glucose (mmol/l)
6.1f	M	+/+	151	5.7	115	13	28	2.38	2.21	83	43	49	59	32	12.8
6.1h	M	+/+	151	5.9	115	12	30	2.45		81	49		57		12.1
7.1e	M	+/+	154	5.3	114	9	29	2.48			96				14.8
8.1k	M	+/+	147	7.5	112	13	32	2.41	2.02						10.0
9.1g	M	+/+	152	5.6	115	11	25	2.42		86	158		62	33	10.3
9.1h	M	+/+	149	7.2	114	10	28	2.42	2.00	54	+		30		8.6
9.1i	M	+/+	153	5.6	115	11	27	2.44			26				11.6
11.1e	M	+/+	157	6.0	118	11	32	2.49		94	+		63	30	7.7
8.1g	M	M/+	154	9.6	121										
8.1j	M	M/+	154	8.4	120										
8.1l	M	M/+	151	6.5	115	11	29	2.39		80	60	57	61	31	11.8
10.1d	M	M/+	154	9.2	119	13	33	2.32		28	145		-		7.2
10.1e	M	M/+	153	8.0	117	13	36	2.38		83	46		53		11.1
6.1a	F	+/+	147	8.5	114	11	32								8.3
6.1e	F	+/+	149	8.5	119	12									10.4
8.1a	F	+/+	151	5.6	117	12			2.60						
9.1a	F	+/+	150	7.2	117	11	30	2.57		127	48				9.2
10.1b	F	+/+	151	6.4	115	10	29	2.51							12.0
11.1a	F	+/+	149	5.3	114	12	31	2.45			40				12.4
11.1c	F	+/+	150	6.1	116	10									12.0
8.1c	F	M/+	149	6.8	115	13	34	2.50	-	117	45	62	60	37	10.4
8.1e	F	M/+	148	7.3	117	13	33	2.45			36				11.1
9.1d	F	M/+	149	7.2	116	12	31								10.2
9.1e	F	M/+	150	8.2	118	12	32								10.4

Metabolic bleed: Test 2																	
ID	Sex	Genotype	Sodium (mmol/l)	Potassium (mmol/l)	Chloride (mmol/l)	Urea (mmol/l)	Creatinine (mol/l)	Tot. Calcium (mmol/l)	Phos (mmol/l)	ALP (U/l)	ALT (U/l)	AST (U/l)	Tot. Protein (g/l)	Albumin (g/l)	Tot. Cholesterol (mmol/l)	Triglycerides (mmol/l)	Glucose (mmol/l)
6.1f	M	+/+	150	7.9	113	11	30	2.46		66	42	59	62	32	2.5	1.80	8.9
6.1h	M	+/+	151	7.2	117	8			2.16								
7.1e	M	+/+	150	6.9	112	8		2.49	2.56	106	+	374	58	28	2.9	1.9	8.5
8.1k	M	+/+	148	7.7	112	12	29	2.46	1.98		48						10.6
9.1g	M	+/+	151	7.3	115	10	22		2.35								7.8
9.1h	M	+/+	151	7.8	116	9	27	2.43		102	179	133	56	29	2.6		9.2
9.1i	M	+/+	149	6.5	112	11	26		2.38								11.0
11.1e	M	+/+	153	7.6	116	10	26	2.54		118	161						7.8
8.1g	M	M/+															
8.1j	M	M/+	151	7.6	116	9	23	2.33	2.38	107	94	111	56	29	2.7	1.4	9.2
8.1l	M	M/+	149	6.2	111	9	27	2.49	1.8	73	45	47	63	34	4.2	3.5	10.5
10.1d	M	M/+	152	7.7	115	9	27	2.49		120	173	100	60	31	2.3		6.6
10.1e	M	M/+	150	7.4	111	10	29	2.49	2.14	65	45	47	64	33	4.5	2.6	9.3
6.1a	F	+/+	146	6.4	114	8	28	2.40	1.98	105	31						10.6
6.1e	F	+/+	144	8.5	111	10	29										9.2
8.1a	F	+/+															
9.1a	F	+/+	149	7.7	114	9	31	2.59	2.12	113	40	60	64	34	2.8	4.4	8.5
10.1b	F	+/+	149	7.1	114	7	26	2.6			42						9.6
11.1a	F	+/+	147	7.1	111	9	31	2.53		112	38	54	62	36	2.9		8.7
11.1c	F	+/+	150	7.1	112	7	30	2.64	1.99	84	43	58	67	37	2.2		9.3
8.1c	F	M/+	146	5.6	111	12	28	2.55	1.85	109	42	52	74	34	2.7	3.7	10.3
8.1e	F	M/+	145	7.7	112	9	29	2.47	2.17		34						8.8
9.1d	F	M/+	147	8.2	113	10	32	2.61			40						7.1
9.1e	F	M/+	145	6.6	114	8	29	2.40		120	31	49	58	34			9.9

Key: + = Value too high to be real therefore excluded from dataset
- = Value too low to be real therefore excluded from dataset

Appendix 5.5
Grip Strength data

Grip Strength																		
Mouse Details				Forelimb Grip Strength (g/cm^3)							Forelimb & hindlimb grip strength (g/cm^3)						
Mouse ID	Genotype	Sex	Date	1	2	3	4	5	Total	Ave	1	2	3	4	5	Total	Ave
10.1d	M / +	M	12/5/04	128	90	119	123	114	574	115	238	221	200	203	233	1095	219
10.1e	M / +	M	12/5/04	130	107	119	119	128	603	121	228	216	197	212	209	1062	212
6.1g	M / +	M	12/5/04	82	87	118	100	106	493	99	212	214	221	164	130	941	188
8.1g	M / +	M	14/5/04	171	126	136	104	110	647	129	238	175	240	214	208	1075	215
8.1h	M / +	M	14/5/04	93	71	100	88	91	443	89	210	195	224	178	228	1035	207
8.1j	M / +	M	14/5/04	77	121	149	128	139	614	123	204	224	195	193	193	1009	202
8.1l	M / +	M	17/5/04	89	93	102	98	87	469	94	197	192	191	198	187	965	193
9.1f	M / +	M	21/5/04	88	130	112	112	89	531	106	212	233	250	215	222	1132	226
6.1f	+ / +	M	12/5/04	65	72	77	88	86	388	78	147	138	144	160	162	751	150
6.1h	+ / +	M	12/5/04	124	127	133	92	156	632	126	234	142	182	183	188	929	186
7.1e	+ / +	M	14/5/04	99	106	102	98	112	517	103	175	184	173	188	192	912	182
8.1k	+ / +	M	17/5/04	95	125	73	89	76	458	92	257	197	193	245	175	1067	213
9.1g	+ / +	M	21/5/04	112	123	145	143	88	611	122	235	254	216	213	198	1116	223
9.1h	+ / +	M	21/5/04	141	105	137	100	141	624	125	239	225	193	196	179	1032	206
9.1i	+ / +	M	21/5/04	120	121	154	99	197	691	138	197	203	224	198	212	1034	207
11.1e	+ / +	M	17/5/04	96	10	114	96	111	427	85	199	158	198	157	187	899	180
10.1a	M / +	F	12/5/04	81	97	121	101	95	495	99	140	162	165	150	198	815	163
6.1c	M / +	F	12/5/04	97	112	91	109	120	529	106	155	143	140	124	148	710	142
6.1d	M / +	F	12/5/04	63	61	134	73	110	441	88	156	145	123	138	158	720	144
8.1b	M / +	F	14/5/04	102	104	94	97	102	499	100	154	157	147	152	145	755	151
8.1c	M / +	F	14/5/04	131	127	141	125	137	661	132	156	145	171	154	162	788	158
8.1d	M / +	F	14/5/04	87	87	120	83	85	462	92	141	131	114	143	119	648	130
9.1d	M / +	F	17/5/04	58	76	91	92	89	406	81	135	133	140	175	195	778	156
9.1e	M / +	F	17/5/04	60	76	78	95	67	376	75	167	157	178	157	146	805	161
10.1b	+ / +	F	12/5/04	136	95	111	115	118	575	115	188	175	156	157	170	846	169
6.1a	+ / +	F	12/5/04	98	94	95	108	115	510	102	139	157	140	156	163	755	151
6.1e	+ / +	F	12/5/04	112	108	105	107	108	540	108	149	144	171	161	154	779	156
8.1a	+ / +	F	14/5/04	90	96	103	96	100	485	97	157	162	171	155	162	807	161
8.1e	+ / +	F	14/5/04	93	70	75	67	75	380	76	164	151	142	152	149	758	152
9.1a	+ / +	F	17/5/04	82	85	89	91	102	449	90	145	158	165	154	154	776	155
11.1a	+ / +	F	17/5/04	101	91	101	112	97	502	100	155	157	165	177	165	819	164
11.1c	+ / +	F	17/5/04	98	97	122	94	102	513	103	191	181	173	195	187	927	185

Appendix 5.6
Spontaneous Alternation data

Foxf2 W174R Phenotyping Screen: Spontaneous Alternation							
Mouse ID	Genotype	Sex	Date	Latency (secs)	Alternations	Total No. arms -2	% Alt
10.1d	M/+	M	14/05/2004	9	17	26	65
10.1e	M/+	M	14/05/2004	17	12	19	63
6.1g	M/+	M	14/05/2004	34	9	11	82
9.1f	M/+	M	24/05/2004	90	5	12	42
8.1g	M/+	M	18/05/2004	13	0	2	0
8.1h	M/+	M	18/05/2004	149	0	4	0
8.1j	M/+	M	18/05/2004	11	11	15	73
8.1l	M/+	M	18/05/2004	18	7	14	50
6.1f	+/+	M	14/05/2004	105	5	9	56
6.1h	+/+	M	14/05/2004	13	9	17	53
7.1e	+/+	M	18/05/2004	300	0	0	0
8.1k	+/+	M	18/05/2004	12	10	19	53
9.1g	+/+	M	24/05/2004	25	21	12	175
9.1h	+/+	M	24/05/2004	12	11	23	48
9.1i	+/+	M	24/05/2004	34	10	15	67
11.1e	+/+	M	18/05/2004	123	4	10	40
6.1c	M/+	F	14/05/2004	17	11	18	61
6.1d	M/+	F	14/05/2004	21	11	21	52
10.1a	M/+	F	14/05/2004	46	9	17	53
8.1b	M/+	F	18/05/2004	47	2	7	29
8.1c	M/+	F	18/05/2004	25	7	15	47
8.1d	M/+	F	18/05/2004	300	0	0	0
9.1d	M/+	F	18/05/2004	8	6	8	75
9.1e	M/+	F	18/05/2004	55	7	14	50
10.1b	+/+	F	14/05/2004	7	5	15	33
6.1a	+/+	F	14/05/2004	52	19	30	63
6.1e	+/+	F	14/05/2004	22	10	17	59
8.1a	+/+	F	18/05/2004	17	14	19	74
8.1e	+/+	F	18/05/2004	106	13	19	68
9.1a	+/+	F	18/05/2004	17	13	23	57
11.1a	+/+	F	18/05/2004	284	0	2	0
11.1c	+/+	F	18/05/2004	257	0	2	0

Appendix 5.7
Intra-peritoneal Glucose Tolerance Screen Data

Foxf2 W174R Phenotyping Screen: IPGTT						
ID	Sex	Transgenic?	Date	T_0	T_{60}	T_{120}
8.1k	m	+/+	25/06/04	6.12	9.4	fail
9.1g	m	+/+	25/06/04	5.93	7.83	5.47
9.1h	m	+/+	25/06/04	7.46	9.64	6.76
10.1d	m	M/+	25/06/04	5.82	7.65	5.54
10.1e	m	M/+	25/06/04	7.7	8.92	3.97
8.1g	m	M/+	25/06/04	7.17	10.25	7.93
6.1a	f	+/+	25/06/04	6.21	6.91	5.19
6.1e	f	+/+	25/06/04	7.04	8.02	6.21
9.1a	f	+/+	25/06/04	4.23	8.35	5.69
9.1d	f	M/+	25/06/04	4.99	7.65	6.13
9.1e	f	M/+	25/06/04	5.62	7.32	4.99
8.1c	f	M/+	25/06/04	4.29	7.76	5.47

Appendix 5.8
Light-Dark Box Behavioural Data

Foxf2 W174R Phenotyping Screen: Light-dark box														
ID	Genotype	Sex	Date	Latency (seconds)	Move episodes	Move time (seconds)	Velocity	Distance	L-Half time (seconds)	R-Half time (seconds)	R-Half Entries	V-plane entries	L-Half Horz CNT	R-Half Horz CNT
8.1g	M/+	M	14/06/04	23	42	246	499.7	99.9	136	164	5	33	18	23
8.1h	M/+	M	14/06/04	19	33	253	759.3	151.8	119	181	10	40	9	19
8.1j	M/+	M	14/06/04	35	48	234	802.1	160.4	107	193	10	60	12	32
10.1d	M/+	M	14/06/04	3	24	257	898.9	179.7	136	164	11	53	8	14
10.1e	M/+	M	14/06/04	5	26	266	923.3	184.6	171	129	14	56	12	11
8.1l	M/+	M	14/06/04	39	39	247	916.2	183.2	124	176	13	53	13	23
6.1f	+/+	M	14/06/04	5	28	259	776	155.2	181	119	6	30	16	11
6.1h	+/+	M	14/06/04	13	30	262	833.2	166.6	143	157	10	37	10	19
7.1e	+/+	M	14/06/04	24	27	263	792.9	158.5	163	137	9	41	14	10
8.1k	+/+	M	14/06/04	9	33	252	903.5	180.7	139	161	13	47	14	16
9.1g	+/+	M	14/06/04	11	32	253	843	168.6	135	165	11	50	10	21
9.1h	+/+	M	14/06/04	48	31	263	979.2	195.8	140	160	10	51	12	16
9.1i	+/+	M	14/06/04	23	31	249	926.5	185.3	109	191	10	56	10	19
11.1e	+/+	M	14/06/04	22	33	259	900.5	180.1	120	180	11	58	11	21
6.1d	M/+	F	14/06/04	18	33	263	1028.9	205.7	108	192	13	41	10	21
8.1c	M/+	F	14/06/04	21	31	261	1004.1	200.8	158	142	12	50	13	14
8.1e	M/+	F	14/06/04	41	34	261	864.7	172.9	147	153	12	48	15	19
9.1d	M/+	F	14/06/04	12	36	246	854.7	170.9	174	126	11	39	19	13
9.1e	M/+	F	14/06/04	37	42	242	774.7	154.9	112	188	7	39	11	26
6.1e	+/+	F	14/06/04	66	38	253	714.7	142.9	178	122	12	31	18	16
8.1a	+/+	F	14/06/04	10	42	251	862.4	172.4	102	198	13	49	9	27
9.1a	+/+	F	14/06/04	28	27	268	994.1	198.8	165	135	12	46	10	10
10.1b	+/+	F	14/06/04	39	43	240	585.9	117.1	192	108	9	28	31	7
11.1a	+/+	F	14/06/04	22	43	239	879.6	175.9	114	186	14	54	10	24
11.1c	+/+	F	14/06/04	12	35	250	881.4	176.2	120	180	15	55	11	21

Appendix 5.9
Pre-Pulse Inhibition Behavioural Data

Foxf2 W174R Phenotyping Screen: Pre-Pulse Inhibition							
ID	Genotype	Sex	Date	Box	dB	% inhibition 12Hz	% inhibition 20Hz
10.1d	M/+	M	17/05/04	1	90	81.16	74.66
10.1e	M/+	M	17/05/04	2	90	70.26	70.18
6.1g	M/+	M	17/05/04	11	90	54.35	64.99
8.1g	M/+	M	19/05/04	7	90	52.87	34.77
8.1h	M/+	M	19/05/04	8	90	61.13	53.48
8.1j	M/+	M	19/05/04	9	90	73.03	71.72
8.1l	M/+	M	19/05/04	12	90	73.49	71.91
9.1f	M/+	M	19/05/04	1	90	73.79	78.77
6.1f	+/+	M	17/05/04	10	90	N.A.	67.89
6.1h	+/+	M	17/05/04	12	90	56.4	52.11
7.1e	+/+	M	19/05/04	12	90	49.94	53.92
8.1k	+/+	M	19/05/04	11	90	61.26	67.67
9.1g	+/+	M	26/05/04	2	90	49.85	54.43
9.1h	+/+	M	26/05/04	3	90	58.68	44.97
9.1i	+/+	M	26/05/04	4	90	56.25	49.3
11.1e	+/+	M	19/05/04	2	90	39.4	49.94
10.1a	M/+	F	17/05/04	3	90	52.18	40.44
6.1c	M/+	F	17/05/04	7	90	70.43	73.3
6.1d	M/+	F	17/05/04	8	90	67.36	60.76
8.1b	M/+	F	19/05/04	2	90	69.63	56.88
8.1c	M/+	F	19/05/04	3	90	69.09	58.1
8.1d	M/+	F	19/05/04	4	90	66.06	48.43
9.1d	M/+	F	19/05/04	7	90	46.15	42.93
9.1e	M/+	F	19/05/04	8	90	35.79	18.88
10.1b	+/+	F	17/05/04	4	90	31.7	22.36
6.1a	+/+	F	17/05/04	5	90	45.79	41.47
6.1e	+/+	F	17/05/04	9	90	60.53	57.26
8.1a	+/+	F	19/05/04	1	90	73.39	75.71
9.1a	+/+	F	19/05/04	3	90	20.22	17.6
8.1e	+/+	F	19/05/04	5	90	58.71	45.46

11.1a	+/+	F	26/05/04	9	90	62.45	60.43
11.1c	+/+	F	19/05/04	11	90	56.12	54.26

PPI Re-Tests

ID	Genotype	PP	P	%Inhibition	PP	P	%Inhibition	PP	P	%Inhibition	PP	P	%Inhibition
Foxf2/8.1j	M/+	70	110	26.30	80	110	64.46	85	110	58.18	90	110	57.92
Foxf2/8.1l	M/+	70	110	20.15	80	110	69.64	85	110	62.85	90	110	46.43
Foxf2/8.1g	M/+	70	110	11.30	80	110	68.34	85	110	74.96	90	110	60.22
foxf2/8.1h	M/+	70	110	26.57	80	110	66.38	85	110	79.53	90	110	61.73
Foxf2/10.1d	M/+	70	110	7.76	80	110	49.22	85	110	62.85	90	110	56.67
Foxf2/10.1e	M/+	70	110	32.14	80	110	64.67	85	110	60.82	90	110	51.58
Foxf2?6.1g	M/+	70	110	5.25	80	110	46.41	85	110	59.76	90	110	47.23
Foxf2/9.1f	M/+	70	110	12.36	80	110	54.71	85	110	66.72	90	110	56.73
AVERAGE				**17.7**			**60.5**			**65.7**			**54.8**
STDV				**9.3**			**9.0**			**7.7**			**5.8**
Foxf2/6.1f	+/+	70	110	1.38	80	110	65.56	85	110	70.67	90	110	55.74
Foxf2/6.1h	+/+	70	110	13.78	80	110	75.11	85	110	70.27	90	110	64.95
Foxf2/11.1e	+/+	70	110	4.65	80	110	71.48	85	110	67.86	90	110	51.71
Foxf2/7.1e	+/+	70	110	26.05	80	110	59.84	85	110	76.73	90	110	70.20
Foxf2/8.1k	+/+	70	110	-12.34	80	110	53.95	85	110	54.94	90	110	50.21
Foxf2/9.1g	+/+	70	110	37.14	80	110	78.42	85	110	79.94	90	110	70.28
Foxf2/9.1h	+/+	70	110	-11.10	80	110	50.34	85	110	41.36	90	110	46.32
Foxf2/9.1i	+/+	70	110	9.06	80	110	59.92	85	110	67.90	90	110	51.68
AVERAGE				**8.58**			**64.33**			**66.21**			**57.64**
STDV				**17.07**			**10.07**			**12.45**			**9.48**

Appendix 5.10
Acoustic Startle Data

Foxf2 W174R Phenotyping Screen: Acoustic Startle							
Screen Information							
Mouse ID	Genotype	Sex	Date	Box	White noise	12 Hz	20 Hz
10.1d	M/+	M	17/05/04	1	677	663	990
10.1e	M/+	M	17/05/04	2	1243	1183	1458
6.1g	M/+	M	17/05/04	11	889	447	992
8.1g	M/+	M	19/05/04	7	529	1071	1076
8.1h	M/+	M	19/05/04	8	1225	1255	1062
8.1j	M/+	M	19/05/04	9	1558.6	1437	1170
8.1l	M/+	M	19/05/04	12	978	848	844.6
9.1f	M/+	M	26/05/04	1	1290	1125	1224
6.1f	+/+	M	17/05/04	10	705	641	631
6.1h	+/+	M	17/05/04	12	1593	1478	1810
7.1e	+/+	M	19/05/04	12	611	966	1247
9.1g	+/+	M	26/05/04	2	1518	1488	1810
9.1h	+/+	M	26/05/04	3	1179	805	853
9.1i	+/+	M	26/05/04	I	1015	1085	1032
11.1e	+/+	M	19/05/04	2	869.2	926	1019
8.1k	+/+	M	19/05/04	11	804.4	861	820
10.1a	M/+	F	17/05/04	3	622	489	990
6.1c	M/+	F	17/05/04	7	855	557	698
6.1d	M/+	F	17/05/04	8	821	998	1116
8.1b	M/+	F	19/05/04	2	726	922	1251
8.1c	M/+	F	19/05/04	3	635	322	793
8.1d	M/+	F	19/05/04	4	493	770	750
9.1d	M/+	F	19/05/04	7	666	378	949
9.1e	M/+	F	19/05/04	8	512	499.2	430.6
10.1b	+/+	F	17/05/04	4	597	551	597
6.1a	+/+	F	17/05/04	5	413	517	663
8.1a	+/+	F	19/05/04	1	738	630	627
8.1e	+/+	F	19/05/04	5	865	668	752
9.1a	+/+	F	19/05/04	3	581.2	1012	660
6.1e	+/+	F	17/05/04	9	920	870	980
11.1a	+/+	F	19/05/04	9	676	933	811
11.1c	+/+	F	19/05/04	11	271	401	325

Appendix 5.11

Correction for Multiple Testing

One possibility requiring consideration is that the observations which indicated a significant difference between mutants and wildtype data are caused by random events. The only significant difference that was observed in the initial test cohort was between mutant and wildtype males at 12 kHz during tests for PPI. These tests were only a small fraction of the total number of tests that these mice were subjected to. The most severe correction for multiple testing, the Bonferroni correction, would multiply the raw p-values by the total number of independent tests [294].
When correcting a dataset with the Bonferroni method, tests that are not independent of the test of interest should not be included because the probability that they would give the same result when testing a null hypothesis is greater than if the data resulted from independent tests. The determination of whether a test is independent however, can be difficult. Some tests could mistakenly be considered as independent when phenotypic links exist but are not immediately apparent. There is also the issue of varying degrees of relatedness between tests and how strong a correlation is required before omitting a test from a Bonferroni correction.

One example of a strong correlation is that deaf mice would show an absence of the preyer reflex and would also not respond to startle. The Startle test is an inherent component of PPI trials, so a healthy hearing range is also required for a normal PPI response.

The replicated tests for males and females are obviously correlated and the numbers of tests taken into account for a Bonferroni correction are therefore not doubled during a correction when both sexes are analysed within separate cohorts.

That anxiety can be tested with acoustic startle is clearly recognised, well documented [251, 252, 254, 295] and should therefore be considered as correlated. The numerous anxiety tests that were performed on the $Foxf2^{W174R}$ mutants therefore should all be omitted from a Bonferroni correction for Startle data. These tests were: Transfer arousal, Locomotive Activity, positional passivity (2 tests), touch escape, and all 9 of the Light/Dark box tests.

The association between PPI and anxiety is not nearly as extensively documented as the relationship between the startle response and anxiety. There is evidence to suggest that patients with panic disorder and post traumatic stress (both of which are regarded as anxiety related conditions) exhibit a reduced PPI [296, 297]. However, the correlation is not nearly as pervasive as it is with startle and this weaker correlation was considered low enough to allow the anxiety tests to remain in the PPI Bonferroni corrections.

Links between PPI and Startle with aggression have been identified in mice [298-300] and humans [301, 302]. There are therefore potential links between the biting/aggression observation with PPI and the acoustic startle test. However, such correlations are unlikely to be frequent enough to interfere with most PPI analysis

and the link was too weak to warrant losing the test from the correction without some evidence of positive correlation within this study.

The majority of the "physical appearance" observations that were required during SHIRPA analysis were not applied to any multiple testing correction. Potential outcomes of most of these observations were only likely to deviate between categories if there was an anomaly in the features of the test animals and therefore, for most of the tests every individual produced the same "normal" result (table 5.2). These data are not normally distributed and the probability of these tests producing a significant result by chance would be very much lower than 0.05%. It therefore did not seem reasonable to include these observations in the correction calculations.

This was also the case for the majority of the nominal behavioural observations which required a mouse frequency as output of qualitative rather than quantitative data (table 5.1), although more of these behavioural tests showed variability than the physical appearance observations.

The method that was used to determine whether or not these categorical tests were included in a correction for multiple testing was the selection of only those tests which demonstrated some variability between the category outputs, including those tests where only one individual deviated from the norm. This is because the groups that exhibited some level of background variation could realistically produce a significant difference by chance.

If this selection procedure is used, the only appearance observation which would contribute towards the correction is the "head shape" record. There were seven frequency based behavioural tests that resulted in the allocation of mice to more than one category; these were: Transfer arousal, Startle Response, Positional Passivity (two observations), Wire manoeuvre, Limb grasping and Contact righting reflex. These methods would be added to the quantitative tests (including the individual metabolic and light/dark box tests) to bring the total number of tests involved in any correction to 65

63 of the 65 phenotyping tests were classified as independent of PPI data, the exceptions being Startle and the preyer reflex observation. This would mean that $p=0.009$ for the original PPI test would be corrected to $p=0.567$ with the Bonferroni adjustment (or the significance threshold would be adjusted to 0.0008)- and therefore the result is not significant. Further PPI tests were not corrected because the only other secondary tests performed were the startle tests which are strongly correlated to PPI.

The Bonferroni adjustment is considered conservative by many statisticians. One alternative is to control the false discovery rate (FDR) [303]. There are many alternative methods for the determination of the FDR threshold, but one that seemed particularly appropriate was a method designed by Benjamini and Lieu which

Benjamini states has no requirement for the elimination of non-independent tests from the calculation [303, 304].

This method requires that the data is first tabulated with the tests ordered by their significance, with the lowest p value first. A threshold is then calculated for each test with the following equation:

$$h = \frac{0.05 \times m/}{(m+1-i)2}$$

Where h = the significance threshold for a specific test
m = the total number of tests requiring correction
i = the rank of the specific test within the table (i.e.: lowest p value is ranked 1, the second lowest = 2 etc).

The threshold for the first test is then compared with it's p value, and if the p value is below the threshold then the result is considered significant and the procedure is repeated for the next test listed on the table. This process continues until the first test with a higher p value than its threshold is discovered. This is classified as having no significant difference between the analysed groups, as are all the subsequent tests that follow in the table.

Within the male Foxf2^{W174R} dataset, the lowest p value was for potassium test 1, where p = 0.02. A total of 65 tests were included in the correction which included two tests for each metabolite for the blood chemistry analysis. The lowest threshold, which is equal to the Bonferroni threshold for all tests, is 0.0008 - well below the lowest p value. Even if only 1 test per metabolite contributes to the correction (to avoid the inclusion of the repeated tests), the threshold remains significantly lower than the lowest observed p value at p=0.0009 for 55 tests. Therefore none of the tests are significant with either correction protocol.

Other significant results in the initial cohort of animals included levels of blood potassium, calcium and chloride in males and blood urea in females. No obvious correlation existed between the potential results of these tests and the other primary screen tests, so each of the significant results were corrected by 65.

The significance adjustments from the two different correction methods are equal for the sample with the lowest p value. There is an increasingly lenient threshold as the p value decreases within the ranked table for the Benjamini and Liu method (Table 5.10a). However in these data, none of the tests were significant.

Another, more commonly utilised method for the determination of the false discovery rate was also designed by Benjamini, alongside a different co-worker, Hochberg [305]. This procedure requires that only independent data is applied to the correction. The startle test and the preyer observation were excluded from the Benjamini & Hochberg correction of PPI significance as with the execution of the

Bonferroni correction. The results from this method also shifted the significance thresholds for each test below their respective p values, meaning that there were no significant results in the corrected data as shown in appendix 5.11.

Test	p value	Rank	BL Threshold
Blood Potassium 1	0.002	1	0.0008
Blood Chloride 1	0.008	2	0.0008
PPI. 12khz@90dB	0.009	3	0.0008
Blood Calcium 1	0.017	4	0.0008
Urine Protein	*0.049	5	0.0009
Urine Ascorbic acid	*0.06	6	0.0009
Blood Creatinine 1	0.068	7	0.0009
PPI. 20khz@90dB	0.070	8	0.0010
Light - Dark Move time	0.140	9	0.0010
Grip strength Fore & Hindlimbs	0.150	10	0.0010
Light - Dark Move episodes	0.190	11	0.0011
PPI retest. WN@70dB	0.210	12	0.0011
Blood AST 2	0.229	13	0.0012
R-Half Horz CNT	0.260	14	0.0012
Urine pH	0.260	15	0.0012
Blood Urea 1	0.263	16	0.0013
Blood Albumin 2	0.263	17	0.0014
Light - Dark Velocity	0.290	18	0.0014
Blood Cholesterol 2	0.297	19	0.0015
Number of pellets	0.330	20	0.0015
Blood Sodium 1	0.343	21	0.0016
Startle/Preyer reflex	0.350	22	0.0017
Blood ALP 1	0.354	23	0.0018
Blood Phosphate 2	0.364	24	0.0018
Spontaneous Alternation Latency	0.390	25	0.0019
PPI retest. WN@80dB	0.430	26	0.0020
Blood Urea 2	0.433	27	0.0021
Blood Protein 2	0.457	28	0.0023
Startle. 20khz@90dB	0.460	29	0.0024
Light - Dark Distance	0.460	30	0.0025
struggles when scruffed	*0.4667	31	0.0027
Light - Dark L-Half time	0.470	32	0.0028
Light - Dark R-Half time	0.470	33	0.0030
PPI retest. WN@90dB	0.480	34	0.0032
Spontaneous Alternation	0.500	35	0.0034
Blood Chloride 2	0.544	36	0.0036
Blood Calcium 2	0.546	37	0.0039
Blood Glucose 1	0.563	38	0.0041
struggles when held by tail	*0.5692	39	0.0500
Contact righting reflex	*0.5692	40	0.0500
Light - Dark V-plane entries	0.600	41	0.0052
Startle. 12khz@90dB	0.610	42	0.0056
Head shape	*0.61	43	0.0500
Weight	0.650	44	0.0067
Blood Potassium 2	0.695	45	0.0074
Locomotive activity	0.700	46	0.0081
Blood ALP 2	0.710	47	0.0090
Blood ALT 2	0.712	48	0.0100
Light - Dark R-Half Entries	0.720	49	0.0112
Blood Protein 1	0.800	50	0.0127
Blood Glucose 2	0.813	51	0.0144
Blood ALT 1	0.820	52	0.0166
Startle. WN@90dB	0.850	53	0.0192
Light - Dark Latency	0.870	54	0.0226
Blood Sodium 2	0.890	55	0.0269
PPI retest. WN@85dB	0.920	56	0.0325
Blood Creatinine 2	0.926	57	0.0401
Light - Dark L-Half Horz CNT	0.940	58	0.0500
Grip Strength: Forelimbs	0.950	59	0.0500
IPGTT	0.970	60	0.0500
Urine Nitrite	*1.000	61	0.0500

Transfer arousal	*1.000	62	0.0500
Wire manoeuvre	*1.000	63	0.0500
Limb Grasping	*1.000	64	0.0500
Urine Ketones	*1.000	65	0.0500

The False Discovery Rate calculated for the male cohort using the Benjamini & Liu method [304]. This procedure incorporates the full range of tests on which $Foxf2^{W174R}$ mice were tested. Benjamini claims that there are no inherent assumptions regarding the independence of tests with this method [303, 304]. This would make the correction ideal for these data in which there is a range of complex relationships between potential test outputs. However an increase in the number of tests increases the lowest p value threshold to the same extent as the Bonferroni method so it is difficult to see how this test is not effected by the independence or otherwise of the tests. This method is less conservative than the Bonferroni correction on average, but the removal of the non-independent test data for the Bonferroni correction causes an increase in the Bonferroni significance threshold (to p=0.0009). This, combined with the fact that none of the p values come close to approaching either threshold means that the Bonferroni threshold is effectively more lenient with these data.
The threshold for the test with the most significant p value will be more conservative in the Benjamini and Liu calculation. Asterixes (*) represent those p values that were inferred by the Fisher Exact test from frequency data. All other p values were from continuous data calculated with either two-tailed student T tests or ANOVA with repeated measures.

Test	p value	Rank	BL Threshold
Blood Urea 1	0.025	1	0.0010
Grip strength Fore & Hindlimbs	0.070	2	0.0010
Blood Creatinine 1	0.071	3	0.0011
Blood AST 2	0.074	4	0.0011
Number of pellets	0.080	5	0.0012
Startle. 20khz@90dB	0.100	6	0.0012
Blood Urea 2	0.160	7	0.0013
Urine Protein	*0.180	8	0.0013
Blood Sodium 2	0.187	9	0.0014
Urine Ascorbic acid	*0.240	10	0.0014
PPI. 12khz@90dB	0.290	11	0.0015
Light - Dark Distance	0.300	12	0.0016
IPGTT	0.300	13	0.0017
Locomotive activity	0.340	14	0.0018
Light - Dark Velocity	0.360	15	0.0019
Light - Dark Move episodes	0.420	16	0.0020
Blood Potassium 1	0.435	17	0.0021
Spontaneous Alternation Latency	0.440	18	0.0022
Blood Sodium 1	0.479	19	0.0023
Light - Dark Move time	0.490	20	0.0025
Urine pH	0.490	21	0.0027
Startle. 12khz@90dB	0.500	22	0.0028
Blood Calcium 2	0.514	23	0.0030
Blood ALT 2	0.554	24	0.0033
struggles when held by tail	*0.569	25	0.0035
Weight	0.570	26	0.0038
Blood Potassium 2	0.632	27	0.0041
Blood Chloride 1	0.644	28	0.0044
Blood Glucose 2	0.680	29	0.0048
Startle. WN@90dB	0.710	30	0.0053
Light - Dark Latency	0.740	31	0.0058
Light - Dark L-Half Horz CNT	0.770	32	0.0064
Grip Strength: Forelimbs	0.780	33	0.0071
Blood Creatinine 2	0.789	34	0.0079
R-Half Horz CNT	0.798	35	0.0088
Light - Dark L-Half time	0.800	36	0.0100
Light - Dark R-Half time	0.800	37	0.0113
Blood Glucose 1	0.833	38	0.0130
Blood Chloride 2	0.861	39	0.0151
PPI. 20khz@90dB	0.910	40	0.0177
Light - Dark V-plane entries	0.940	41	0.0211
Spontaneous Alternation	0.980	42	0.0255
Head shape	*1.000	43	0.0315
Wire manoeuvre	*1.000	44	0.0398
Limb Grasping	*1.000	45	0.0500
Urine Ketones	*1.000	46	0.0500
Transfer arousal	*1.000	47	0.0500
Startle/Preyer reflex	*1.000	48	0.0500
struggles when scruffed	*1.000	49	0.0500
Contact righting reflex	*1.000	50	0.0500
Urine Nitrite	*1.000	51	0.0500

Table 5.10b: The False Discovery Rate calculated for the female cohort using the Benjamini & Liu method [304]. * = tests for which the p value was determined from frequency data - with the Fisher Exact test. Other tests using cumulative data were performed with either ANOVA & repeated measures or two-tailed T-tests. As was seen in the male cohort, no tests showed a significant difference between mutants and wildtypes following correction for multiple testing.

The False Discovery Rate for the significant PPI result using the method described by Benjamini & Hochberg [305].

All tests that were not independent of PPI at 12kHz/90dB were removed from the calculation of the Benjamini & Hochberg correction because tests are assumed to have no correlation with the measure to be corrected [303]. The tests that were removed from the correction were all other PPI tests, all startle tests and the preyer reflex test.

Test	p value	Rank	BH Threshold
Blood Potassium 1	0.002	1	0.0009
Blood Chloride 1	0.008	2	0.0009
PPI. 12khz@90dB	0.009	3	0.0010
Blood Calcium 1	0.017	4	0.0010
Urine Protein	*0.049	5	0.0010
Urine Ascorbic acid	*0.060	6	0.0011
Blood Creatinine 1	0.068	7	0.0011
Light – Dark Move time	0.140	8	0.0012
Grip strength Fore & Hindlimbs	0.150	9	0.0012
Light – Dark Move episodes	0.190	10	0.0013
Blood AST 2	0.229	11	0.0013
R-Half Horz CNT	0.260	12	0.0014
Urine pH	0.260	13	0.0014
Blood Urea 1	0.263	14	0.0015
Blood Albumin 2	0.263	15	0.0016
Light – Dark Velocity	0.290	16	0.0017
Blood Cholesterol 2	0.297	17	0.0018
Number of pellets	0.330	18	0.0018
Blood Sodium 1	0.343	19	0.0019
Blood ALP 1	0.354	20	0.0020
Blood Phosphate 2	0.364	21	0.0022
Spontaneous Alternation Latency	0.390	22	0.0023
Blood Urea 2	0.433	23	0.0024
Blood Protein 2	0.457	24	0.0026
Light – Dark Distance	0.460	25	0.0027
struggles when scruffed	*0.4667	26	0.0029
Light – Dark L-Half time	0.470	27	0.0031
Light – Dark R-Half time	0.470	28	0.0033
Spontaneous Alternation	0.500	29	0.0036
Blood Chloride 2	0.544	30	0.0038
Blood Calcium 2	0.546	31	0.0041
Blood Glucose 1	0.563	32	0.0045
struggles when held by tail	*0.5692	33	0.0049
Contact righting reflex	*0.5692	34	0.0053
Light – Dark V-plane entries	0.600	35	0.0058
Head shape	*0.61	36	0.0063
Weight	0.650	37	0.0070
Blood Potassium 2	0.695	38	0.0078
Locomotive activity	0.700	39	0.0086
Blood ALP 2	0.710	40	0.0097
Blood ALT 2	0.712	41	0.0109
Light – Dark R-Half Entries	0.720	42	0.0124
Blood Protein 1	0.800	43	0.0143
Blood Glucose 2	0.813	44	0.0166
Blood ALT 1	0.820	45	0.0194
Light – Dark Latency	0.870	46	0.0231
Blood Sodium 2	0.890	47	0.0280
Blood Creatinine 2	0.926	48	0.0346
Light – Dark L-Half Horz CNT	0.940	49	0.0438
Grip Strength: Forelimbs	0.950	50	0.0500
IPGTT	0.970	51	0.0500
Urine Nitrite	*1.000	52	0.0500
Transfer arousal	*1.000	53	0.0500
Wire manoeuvre	*1.000	54	0.0500
Limb Grasping	*1.000	55	0.0500
Urine Ketones	*1.000	56	0.0500

The False Discovery Rate as calculated by the Benjamini &Hochberg method [305]. Tests that correlated with PPI at 12kHz/90dB were removed from the calculation of the Benjamini & Hochberg correction because tests are assumed to have no correlation with the measure to be corrected [303].

The tests that were removed from the correction were all other PPI tests, all startle tests and the preyer reflex test.

Appendices for Chapter 6

Appendix 6.1

Standard Curve 1: RMSEAP1 and RMSEAP2

Transfection Optimisation of EGFP with Genejuice

Genejuice to DNA Ratio	View Field	Total Cell Count (DAPI)	Transfected Cell Count (EGFP)	% Transfected Cells	% Transfected Cells Average
	a	142	19	13.4	
3ul genejuice:	b	128	22	17.2	15.3
1ug DNA	c	-	-	-	
	a	137	13	9.5	
6ul genejuice:	b	84	8	9.5	11.2
1ug DNA	c	136	20	14.7	
	a	106	24	22.6	
3ul genejuice:	b	119	15	12.6	17.6
2ug DNA	c	-	-	-	

Appendix 6.2

SEAP Concentration (pg/ml)	Transfection A: RLU (Relative Light Units)	Transfection B: RLU (Relative Light Units)
4000000	310.7	390.3
2000000	179.2	278
1000000	107.8	186.1
500000	59.34	114.2
100000	12.37	27.82
20000	2.704	6.377
4000	0.5313	1.661
800	0.1248	0.5489
160	0.1035	0.335
0	-	0.273

Appendix 6.3

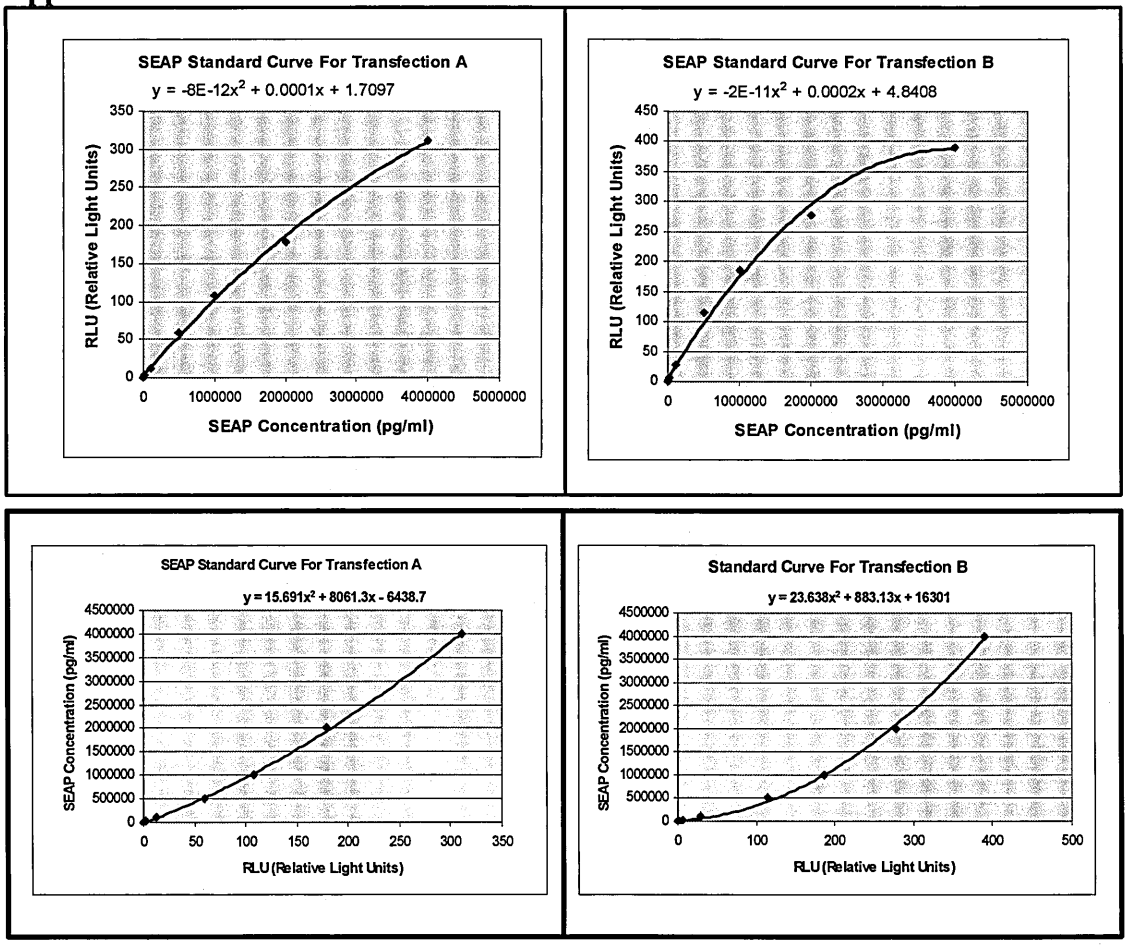


Figure 6.26: Standard SEAP Curves for the two transfections of the FOXF2/pEVRF0 constructs that were co-transfected using pTAL-SEAP as a transfection control. Each data set was entered into an excel table, from which a polynomial trend-line was added. The top two graphs represent the data in the usual format for a standard curve, the bottom two graphs are in a format which enabled the generated equation to be used to calculate SEAP concentrations from RLU output. The trend-line equation for transfection A was: y = 15.691x2 + 8061.3x - 6438.7. The trend-line equation for transfection B was y = 23.638x2 + 883.13x + 16301.

Appendix 6.4

Cell Counts of DAPI Stained Cells and EGFP Transfected Cells										
Reprter Vector	Transfection	View Field	Empty pEVRF0		V410F FOXF2		W174R FOXF2		Wildtype FOXF2	
			DAPI	EGFP	DAPI	EGFP	DAPI	EGFP	DAPI	EGFP
pTAL-SEAP & FOXF2 Binding Site	1	a	347	53	227	53	218	91	262	51
		b	309	66	188	47	372	74	133	18
		c	282	69	240	48	265	63	116	17
		d							129	19
		e							136	34
	2	a	331	67	101	12	242	76	229	40
		b	304	69	114	17	299	83	144	24
		c	271	49	123	32	246	67	137	20
		d			129	21			155	21
		e			138	16			130	11
	3	a	371	86	114	25	207	74	161	35
		b	289	60	111	23	266	60	153	9
		c	278	63	97	12	340	65	135	23
		d			185	23			130	26
		e			165	25				
Empty pTAL-SEAP	1	a	281	32	319	60	264	54	297	30
		b	321	43	335	60	324	69	242	38
		c	256	37	282	51	246	59	323	40
	2	a	276	49	213	44	261	49	350	33
		b	267	42	317	45	298	50	196	51
		c	294	74	319	39	284	73	250	33
	3	a	304	45	285	43	243	63	264	39
		b	289	50	219	43	288	54	305	46
		c	147	56	195	51	212	44	214	50

Cell counts under different lighting conditions which allow for a total cell count because each mount was DAPI-stained, and a count of the EGFP transfected cells. The percentage of cells which had been transfected by EGFP provided a ratio with which to normalise the SEAP data. Cell counts in black were taken on a Zeiss Axioplan 2 Microscope with the associated imaging software Axiovision v4.5. Cell counts in green text were performed on an Olympus BX40 Microscope . Green and blue wavelengths were filtered electronically from this microscope with Smartcapture (Digitalscientific UK Ltd). This microscope had a smaller field of view at identical magnification, so a larger number of view fields were counted to compensate for this.

Appendix 6.5

pEVRF0 Construct	SEAP Production after Transfection A (pg/ml)	SEAP Production after Transfection B (pg/ml)
0	1.06×10^5	4.75×10^4
WR	9.58×10^4	4.82×10^4
VF	6.18×10^6	1.57×10^6
WT	1.29×10^6	7.05×10^5
PG	6.08×10^6	4.29×10^5
NT	-1.82×10^3	1.67×10^4

EMSA SEAP data.

Appendix 6.6

Transactivation Data for Each Sample at 24hrs Post-Transfection								
Sample	No.	RLU	SEAP pg/ml	Normalisation figure	Normalised SEAP/GFP	Average	Std Deviation	Std error
WT & 0	1	18.74	317528.2	12.7	24945.0			
	2	18.98	321659.0	16.2	19835.5			
	3	20.1	340510.3	17.7	19194.8	21325.1	3151.3	1819.4
WT & BS	1	12.1	205278.7	17.5	11745.9			
	2	12.1	205933.2	14.1	14555.2			
	3	10.6	181201.7	16.2	11209.8	12503.6	1796.8	1037.4
W-R & 0	1	9.4	162313.5	21.9	7407.7			
	2	9.9	170649.5	20.4	8357.4			
	3	9.7	167706.3	21.8	7689.3	7818.2	487.8	281.6
W-R & BS	1	5.3	97302.6	28.5	3417.8			
	2	5.6	102724.7	28.8	3566.8			
	3	5.7	104172.1	25.8	4036.5	3673.7	322.9	186.4
V-F & 0	1	14.4	243238.9	18.3	13314.9			
	2	16.8	284910.5	15.7	18155.4			
	3	17.1	290173.5	20.3	14299.8	15256.7	2558.2	1477.0
V-F & BS	1	10.2	175063.9	22.8	7684.1			
	2	10.5	179584.7	16.1	11129.0			
	3	11.1	189954.8	16.5	11497.7	10103.6	2103.5	1214.4
eP & 0	1	5.3	98000.9	13.1	7493.1			
	2	6.7	119254.7	19.6	6099.6			
	3	7.2	126705.8	23.4	5414.9	6335.8	1059.1	611.4
eP & BS	1	4.9	90859.5	20.4	4461.1			
	2	6.5	116573.5	20.3	5731.2			
	3	6.2	111534.7	22.2	5023.8	5072.0	636.4	367.4

The Mean percentage of EGFP transfected cells within each Transactivation assay. These figures were used to normalise SEAP levels. EGFP was the co-transfectant control which was used to provide information regarding the relative transfection efficiency in each sample. DAPI stained cells were counted to give the total cell count and this was compared with the number of green cells that were associated with EGFP transfection.

Appendix 6.7

Transactivation Data for Each Sample at 48hrs Post-Transfection								
Sample	No.	RLU	SEAP pg/ml	Normalisation figure	Normalised SEAP/GFP	Average	Std Deviation	Std Error
WT & 0	1	44.3	720532.8	12.7	56605.0			
	2	40.34	651697.5	16.2	40187.7			
	3	47.7	780004.2	17.7	43969.3	46920.7	8597.4	4963.7
WT & BS	1	25.4	399844.9	17.5	22878.8			
	2	25.7	404932.3	14.1	28620.2			
	3	26.5	417922.2	16.2	25854.1	25784.4	2871.3	1657.8
W-R & 0	1	30.4	483246.6	21.9	22054.4			
	2	31.8	506003.7	20.4	24781.1			
	3	31.1	494946.5	21.8	22693.3	23176.3	1426.1	823.3
W-R & BS	1	27.6	436069.6	28.5	15317.0			
	2	20.9	327267.2	28.8	11363.5			
	3	21.5	337125.8	25.8	13063.0	13247.8	1983.2	1145.0
V-F & 0	1	35.0	560820.2	18.3	30699.4			
	2	38.6	622253.8	15.7	39652.0			
	3	40.2	649453.1	20.3	32005.2	34118.9	4836.1	2792.1
V-F & BS	1	22.7	355770.2	22.8	15615.8			
	2	24.1	378571.1	16.1	23460.4			
	3	24.0	376612.2	16.5	22795.8	20624.0	4349.9	2511.4
eP & 0	1	18.0	280853.0	13.1	21473.8			
	2	20.4	319524.8	19.6	16342.9			
	3	19.9	310830.5	23.4	13283.6	17033.4	4138.5	2389.4
eP & BS	1	16.5	255682.1	20.4	12553.7			
	2	23.3	366505.0	20.3	18018.9			
	3	22.4	351385.9	22.2	15827.3	15466.6	2750.4	1587.9

Appendix 6.8

Transactivation Data for Each Sample at 72hrs Post-Transfection								
Sample	No.	RLU (X3)	SEAP pg/ml	Normalisation figure	Normalised SEAP/GFP	Average	Std Deviation	Std Error
WT & 0	1	132.9	3659267.1	12.7	287471.9			
	2	121	3389129.5	16.2	208994.9			
	3	143.0	3880460.7	17.7	218744.1	238403.6	42773.0	137642.4
WT & BS	1	76.1	2261614.4	17.5	129407.6			
	2	77.1	2286695.4	14.1	161621.3			
	3	79.4	2350283.2	16.2	145396.7	145475.2	16107.0	83990.2
W-R & 0	1	91.3	2660372.1	21.9	121414.3			
	2	95.3	2764683.1	20.4	135398.3			
	3	93.4	2714236.2	21.8	124447.9	127086.8	7356.0	73373.6
W-R & BS	1	82.7	2438037.1	28.5	85636.1			
	2	62.8	1892728.8	28.8	65719.8			
	3	64.6	1944069.6	25.8	75329.3	75561.7	9960.2	43625.6
V-F & 0	1	105.1	3008292.1	18.3	164674.5			
	2	115.9	3268814.0	15.7	208299.5			
	3	120.6	3380060.3	20.3	166570.3	179848.1	24657.9	103835.4
V-F & BS	1	68.0	2040094.4	22.8	89545.9			
	2	72.2	2155644.6	16.1	133587.3			
	3	71.9	2145798.3	16.5	129882.1	117671.8	24428.1	67937.8
eP & 0	1	54.1	1645667.9	13.1	125826.3			
	2	61.3	1852131.8	19.6	94731.7			
	3	59.7	1806251.9	23.4	77191.5	99249.8	24630.2	57301.9
eP & BS	1	49.4	1507926.8	20.4	74037.6			
	2	70.0	2094752.7	20.3	102986.4			
	3	67.2	2017639.2	22.2	90879.6	89301.2	14538.8	51558.1

Full Dataset for the transactivation Assay. The RLU at 75 hours are multiplied by 3 because samples were diluted 3X to stay within the optimum SEAP concentrations for measurement reliability.

Appendix 6.9

Time after Transfection	Transfected Plasmids	WR & 0	VF & 0	EP & 0	WT & BS	WR & BS	VF & BS	EP & BS
24 hrs	WT & 0	12.98	5.83	14.40	8.48	16.96	10.78	15.62
	WR & 0		7.15	1.42	4.50	3.98	2.20	2.64
	VF & 0			8.57	2.65	11.13	4.95	9.79
	EP & 0				5.93	2.56	3.62	1.21
	WT & BS					8.48	2.31	7.14
	WR & BS						6.18	1.34
	VF & BS							4.83
48 hrs	WT & 0	9.34	5.04	11.76	8.31	13.24	10.34	12.37
	WR & 0		4.30	2.42	1.03	3.91	1.00	3.03
	VF & 0			6.72	3.28	8.21	5.31	7.34
	EP & 0				3.44	1.49	1.41	0.62
	WT & BS					4.93	2.03	4.06
	WR & BS						2.90	0.87
	VF & BS							2.03
72 hrs	WT & 0	8.35	4.39	10.44	6.97	12.22	9.06	11.18
	WR & 0		3.96	2.09	1.38	3.87	0.71	2.83
	VF & 0			6.05	2.58	7.82	4.66	6.79
	EP & 0				3.47	1.78	1.38	0.75
	WT & BS					5.24	2.09	4.21
	WR & BS						3.16	1.03
	VF & BS							2.13

Tukey test for the two-factor ANOVA with replication of normalised SEAP data from three time points (24, 48 and 72 hours). Each data-point is a q-value determined with the equation shown below, as described at http://web.mst.edu/~psyworld/tukeyssteps.htm:

$$\frac{M_1 - M_2}{\sqrt{MS_w\left(\frac{1}{n}\right)}}$$

M_1 = the larger of the treatment/group means. M_2 = the smaller of the two means. n = number per treatment/group. MS_w = Mean square within groups.

The two reporters and four expression plasmids meant that 16 degrees of freedom were applied to this comparison. This was used to determine the significance

threshold (determined from the Tukey critical value table at http://web.mst.edu/~psyworld/virtualstat/tukeys/criticaltable.html). The q thresholds were 4.9 for p=0.05 and 6.08 for p=0.01. q values have an inverse relationship with the p value and therefore significance increases with a higher number. Red figures represent all q values above 4.9 and therefore $p<0.05$.

Appendix 6.10

Transfection Parameter	Transactivation Assay		EMSA	
	Transactivation Assay	Transactivation Assay (per cm2)	EMSA	EMSA (per cm^2)
Surface area (cm2)	9.60	1.00	175	1.00
Expression vector mass (ng)	450.00	46.88	15750	90.00
SEAP vector mass (ng)	450.00	46.88	1750	10.00
GFP vector mass (ng)	100.00	10.42	NA	NA
Volume Genejuice (ul)	3.00	0.31	56	0.32
Serum free media (ul)	97.00	10.10	2200	12.57
Total volume (ul)	2104.00	219.17	45126	257.86
Total cells	175000.00	18229.17	5250000	30000.00

Comparison of relative constituent ratios in the transfections for the transactivation assay and the EMSA. The conditions were not to scale because different sources were used to determine initial conditions for each experiment. The amounts of genejuice and serum free media were added to each transfection at similar levels. The pTAL-SEAP vector was present in the transactivation assay at ~5 times the level used in the EMSA. GFP was not added to the EMSA. Cells were seeded at approximately double the density in the EMSA. The expression vector was also added at approximately double the volume per cm^2.